Agricultural Information Worldwide

AN INTERNATIONAL JOURNAL FOR INFORMATION SPECIALISTS IN AGRICULTURE, NATURAL RESOURCES, AND THE ENVIRONMENT

ISSN: 2075-7751

Vol. 4, No. 1, 2011

Agricultural Information Worldwide:

An International Journal for Information Specialists in Agriculture, Natural Resources, and the Environment

Official Journal of the International Association of Agricultural Information Specialists (IAALD)

ISSN: 1998-0027 IAALD Secretariat, P.O. Box 63, Lexington, KY 40588-0063, USA e-ISSN: 2075-7751

Guest Editor: Antoinette Paris Greider, Director, International Programs for University Libraries, University of Kentucky, Lexington, KY, USA. E-mail: Toni.Greider@uky.edu

Assistant Editor: Amélie E.M. Charron, University of Kentucky, Lexington, KY, USA.

Agricultural Information Worldwide is abstracted/indexed in AGRICOLA; CAB Abstracts; Library Literature & Information Science; and Library, Information Science & Technology Abstracts (LISTA).

Membership/Subscriptions

Free to IAALD Members. For current membership/subscription rates, see the IAALD website (http://www.iaald.org). Subscription requests and applications for membership should be directed to: IAALD Subscriptions, c/o IAALD Secretariat, A.P. Greider, Secretary/Treasurer, P.O. Box 63, Lexington, KY 40588-0063, USA, or to toni.greider@iaald.org.

Agricultural Information Worldwide:
An International Journal
for Information Specialists in
Agriculture, Natural Resources,
and the Environment

L'Information Agricole Mondiale:
une Revue Internationale pour
les Spécialistes d'Information sur
l'Agriculture, les Ressources Naturelles
et l'Environnement

Información Agrícola a Nivel Mundial:
Una Revista Internacional para
Especialistas de Información en
Agricultura, Recursos Naturales
y Medio Ambiente

Agricultural Information Worldwide

Vol. 4, no. 1 ▪ 2011

Guest editor: Antoinette Paris Greider

Assistant editor: Amélie E.M. Charron

▪ Official Journal of the International Association of Agricultural Information Specialists (IAALD) ▪

IAALD Secretariat, P.O. Box 63, Lexington, KY 40588-0063, USA

Instructions for Authors

Agricultural Information Worldwide: An International Journal for Information Specialists in Agriculture, Natural Resources, and the Environment (*AIW*) is the official journal of the International Association of Agricultural Information Specialists (IAALD). *AIW* provides an international forum for high quality articles on information, knowledge and communication activities related to the applied life sciences, including agriculture, food from production to marketing, natural resources, fish and wildlife, environment, and agricultural extension and education. Priority will be given to practical and applied topics, such as but not limited to best practices. Research articles with practical applications will also be considered for publication.

Articles submitted will go through a blind review process with an independent reviewer and will be returned to the author for corrections and modifications if necessary. Research should be statistically valid and replicable with the results of broad applicability. English, French, and Spanish language articles will be considered for publication. Generally, full articles should not exceed 5000 words, but longer articles will be considered on a case-by-case basis.

All *AIW* articles are published with a specific tabular style and follow bibliographic conventions as listed in the *Chicago Manual of Style* 15th edition. References should be complete and tables should comply with the editorial style represented in *AIW.* Notes and references should be presented at the end of an article, not as footnotes. An English language abstract of 150 words or less is required at the time of submission. Additional abstracts in French and Spanish are welcome. Articles submitted should be accompanied with the institutional affiliation and address of each author as well as a brief biography.

In addition to full articles, *AIW* also publishes short reports and updates on projects, tools, and organizations in its *AgInfo Dispatches* section. Dispatches will be less formal in nature and will be reviewed for acceptance by the Editor. Dispatch submissions do not require abstracts and should not exceed 1500 words.

AIW is an e-only journal (http://journals.sfu.ca/iaald/index.php/aginfo) housed on a server at Simon Fraser University and articles must be submitted through the system. Contact the editor at info@iaald.org for instructions on submitting an article.

Acknowledgements

The Association would like to thank the following people for contributing to the publication of this issue of *Agricultural Information Worldwide:*

Copy editing – A. Charron

French translations – Marie Josée Jehl-Cooke

Spanish translations – L. Menendez, CIAT

Typesetting/Composition – The Typewright, Lexington, Kentucky, USA

From the Editor's Desk: CIARD and the Interoperability of Agricultural Information

Editorial Acknowledgement: The editor wishes to thank Pamela André, Elizabeth Goldberg, Jeanne Pfander, and Lutishoor Salisbury for their assistance in editing the articles for this issue.

> *IAALD looks forward to a closer co-ordination of its work with the scientist using agricultural publications and with documentalists throughout the world. This is a particularly important field since information in the agricultural area is of utmost importance to world development.*
>
> — Foster Mohrhardt (Mohrhardt, 1962, 135)

As I was gathering background information to prepare to write this editorial, it occurred to me that we have not come very far in our quest to make agricultural information freely available. Foster Mohrhardt struggled with standards and sought to bring the world's agricultural information professionals together to bring some "interoperability" to agricultural information. In our generation we have seen the physical and in some cases the language barriers melt away but we still have one barrier remaining — a general agreement on standards to allow for the exchange of our information. I entered the profession not long after Foster Mohrhardt was ending his career and faced many of the same challenges he faced in organizing and providing agricultural information to all who needed it. When I arrived in the Agricultural Library at the University of Kentucky, about one third of the print collection consisted of Agricultural Experiment Station publications from all the states in the United States and from many stations abroad as well as publications from FAO and other international centers around the world. There was a great deal of information in the volumes but it was very difficult to pull out. The *Office of the Experiment Station Record,* the *Bibliography of Agriculture* and the new database AGRICOLA provided access to the experiment station publications but indexing for these publications became less readily available as the U.S. National Agricultural Library (NAL) struggled to keep up with the indexing. *CAB Abstracts* (later CABI) had a long track record and did an excellent job of describing the research in their abstracting journals. In some cases this surrogate had to be enough as the document was difficult to retrieve. The *FAO Documentation Index*, *Special Indexes* and AGRIS provided access to much of the international "grey" agricultural literature but providing the documents was difficult. Over the years our indexing improved but the issue of getting the document remained. FAO made a valiant attempt by providing their documents through subscription on microfiche. This collection was rich but fairly expensive for most libraries and required some maintenance on the local level. NAL responded by organizing a cooperative microfilming project of all state experiment station publications in the United States. This multi-year project required buy-in from all states in the United States and not all states participated. By the 1980s most of the experiment station publications in the U.S. and soon after abroad ceased (Kentucky's last bulletin was published in 1985) and agricultural research became "legitimized" by being published in the commercial journals. Libraries now had to buy the information which meant if a library could not buy it, it was not readily available. This was a bitter pill to swallow for those of us who believed in the free flow of agricultural information and widened the gap between north and south.

Being good librarians we went to work to find ways to cooperate and aid in the providing of agricultural information. As the commercial journals proliferated AGRICOLA, CABI, and AGRIS (the big three in Agriculture)

Who was Foster Mohrhardt?

Foster Mohrhardt was both a librarian and a documentalist who was one of the founding fathers of IAALD. He served as IAALD's president for the first 15 years of the organization (1955–1969).

> . . . his personal goal [was] to develop national and international library networks to support scientific communication . . . and worked actively throughout his careers to bring librarians and documentalists together to address information problems outlined by practicing scientists and policy-makers . . . (Cragin, 2004, 833)

His professional service includes:

- President of the International Association of Agricultural Librarians and Documentalists (IAALD)
- Vice-President of the American Association for the Advancement of Science (AAAS – publishers of the journal *Science*).
- Founder and Chair of Section T (Information Science)
- President of the Association of Research Libraries (United States)
- Vice-President of International Federation of Library Associations (IFLA)
- President of the Council on Biological Sciences Information, National Academy of Sciences (United States)
- President of the American Library Association

He advocated cooperation and collaboration among librarians and documentalists "to bring the new information retrieval (IR) techniques into the library and to make them more effective in meeting the needs of users seeking information." (Cragin, 2004, 848)

worked hard to provide better access through the indexing. Librarians were concerned that we had to look in too many indexes for information so the services each expanded their coverage and provided their own vocabulary to provide some uniformity in accessing the information in their databases. Then mediated literature searching spread to many libraries and librarians were unhappy with the duplication since the service required that you pay per citation. Database providers offered "deduping" of references which helped but with the advent of CD-ROM we were back to wanting to have to buy only one product to serve all needs. This was a pretty tall order given the nature of agricultural information.

> *"All too often prior investigation and research work are ignored or unused as a result of the inability of the research worker to readily locate and obtain the publications he needs. This is one of the greatest challenges that faces us in the library profession today, the urgent need for local, state, and national cooperative action in collecting, organizing, and providing ready service on all important publications."*
>
> — Foster Mohrhardt
> (Mohrhardt, lecture, 1967)

One area that has remained constant over the years has been the call for standards. The U.S. Experiment Station publications were all published in similar formats (6 x 9 inches in size with the Bulletins, Circulars and Annual Reports from each station bound together). Librarians were not amused when a publication was issued in another size because this required a change to the height of the shelf. Standards have always been a part of our profession and became even more important with the advent of MARC and online catalogs. IAALD has been a leader in the standards discussion including the Universal Agricultural Thesaurus (1989) in an attempt to provide a universal searching language for agriculture. (All parties could not come to an agreement.)

The explosion of the Internet and the World Wide Web as a provider of information provided many opportunities along with the challenges. Authors no longer had "be" published but could easily self-publish without the rigorous review process. No self-respecting scientist wanted his/her publication out on the free web and so commercial journals continued to thrive and began using the web as a vehicle to deliver their publications at a higher cost to libraries. In order to combat this development the Soros Foundation provided US$3million to create the "Budapest Open Access Initiative" (BOAI) in February of 2002. The funding was to provide free access to refereed articles on the Internet. As of 2011, 5412 individuals and 541 organizations have signed on to the declaration.

Without a doubt today's web technology provides searching capabilities we never dreamed of and access to materials that were only available in great libraries. It also provided its challenges with simple searches retrieving millions of documents and no way to do comprehensive searching on anything. There is a good deal of agricultural information on the web but much of it is difficult to retrieve. Distributed services such as AgNIC (Agricultural Networked Information Center) in the United States provide for some cooperative organization of the web. This is just a drop in the bucket compared to what is out there. As a profession we have always been aware of the materials that we knew existed but had no good way to access. (We fondly call it "grey" literature.) This has not changed and will not change in the near future.

We are now entering an age where institutions are taking back their publications and self-publishing. The institutional repository movement is allowing producers to distribute their information much as they did in the days of publication exchange. Now the technology allows the producer to put the document up on a server and allow the publications to be freely distributed. Authors are encouraged to work with commercial publishers to allow them to put their edited articles into their institutional repository so that it can be freely shared around the world. It is because of this movement that a program such as CIARD becomes an imperative and not an option. If we are to realize the dream of access to agricultural information to all who need it, good standards must be adhered to. This is not a new concept. Nearly 20 years ago the Joint Information Systems Committee (JISC) was established in the United Kingdom, "its task was to look at the interoperability…and to describe their resources in order to facilitate identification by end-users." (Schöpfel and Farace, 2009, 1120)

Interoperability is the key term here. We can all develop repositories but we cannot "exchange" our publications if we are not interoperable. To be interoperable we must be developing under a common set of standards. Much like OCLC forced us to change the way we cataloged books, the need for interoperability will cause us to think about how we set up our repositories.

The next challenge for us as professionals will be the semantic web which will bring together objects by using a Universal Resource Identifier (URI) and ontologies (formalisms which define the concepts and relationships in particular applications). This next phase of the web will "link data, allowing querying, integration, and sharing of data from distributed sources in heterogeneous formats, using ontologies to prove an associated and explicit semantic interpretation" (O'Hara and Hall, 2009, 4663). "Grey" literature as we know it will cease to exist as the capabilities of the semantic web will link all the formats together in a single search.

IAALD members are well positioned to meet the challenge of the semantic web. As librarians, we are masters of collecting and organizing and as ICT specialists (docu-

mentalists) we know what to use to get the information into the hands of the users. Together, as agricultural information specialists we can realize the CIARD vision but only if we are willing to work together.

So it is fitting that IAALD, an organization founded on the principles of accessible agricultural information should be one of the founding partners of Coherence in Agricultural Research and Development (CIARD) (launched at the XIIth IAALD World Congress in Japan in 2008) and that we devote this issue of *Agricultural Information Worldwide* to the different types of repositories being developed and that we call on all of our members both librarians (information specialists) and documentalists (ICT professionals) to join together and agree to adhere to standards as we develop our repositories. Our repositories must be interoperable and to be interoperable we must adhere to standards. As a profession, both groups understand the importance of standards and their application very well so adhering to them should not be an issue. Agricultural information professionals in both traditional libraries and in communications technology have led the way in information innovation for a number of years. Now is not the time to lose the way but to step up to the plate and lead the way in applying standards to the development of our repositories.

It is fitting that the first article in this issue is a tribute to one of our own "documentalists", Dr. Anton Mangstl, who has the vision and drive to take information to the user. The remarkable transformation of FAO's information services in a short time frame is the result of his dedication to the end result of what we do as professionals. The remainder of the articles is an interesting mix of what is happening with information management around the world. The first article deals with the transformation of a traditional bibliographic database (AGRIS) into a 21st century information service. The following three articles are examples of the work being done to create both research and institutional repositories and the services to access the contents. The next two articles discuss how CIARD can aid in integrating these collections of documents into a global agricultural information system and case studies on how the concept can be implemented at the country level. The final article provides a way all of our institutions can be a part of this important movement by embracing the CIARD vision and registering in the CIARD Ring.

CIARD: A Global Initiative to Make Agricultural Knowledge Accessible

Coherence in Information for Agricultural Research for Development (CIARD) is an international initiative working to make agricultural research information publicly available and accessible to all. Among its actions are advocating and promoting open access, improving applicability and enabling effective use of data and information in agricultural research and innovation. Widely supported by partner organizations around the globe, it is a growing phenomenon that has gained the attention and involvement of hundreds of information providers. These providers subscribe to the goal of building interoperable and openly accessible information resources and related technical services based on innovative and common policies and practices. CIARD partners are developing new virtual facilities and resources to assist providers of all types to participate in the effort. I urge all IAALD members to learn more about CIARD by visiting the website and registering your information services at http://ring.ciard.net/. Together we can realize the CIARD vision: "To make public domain agricultural research information and knowledge truly accessible to all."

— Barbara Hutchinson, IAALD President 2010–2011

This issue clearly illustrates that the past four decades have been ones of rapid change and achievement but I personally believe that the best is yet to come!

— Antoinette Paris Greider

References

"Budapest Open Access Initiative". http://www.soros.org/openaccess/read.shtml [accessed February, 28, 2011].

"Coherence in Information for Agricultural Research for Development." http://www.ciard.net/ [accessed March 1, 2011]

Cragin, Melissa H. 2004. "Foster Mohrhardt: Connecting the Traditional World of Libraries and the Emerging World of Information Science." *Library Trends*, 52(4):833–852.

"International Co-ordination of Documentation". 1939. *Nature*, 144(3647): 559–560.

Kohl, David F. 2009. "Collection Development in the ARL Library". *Encyclopedia of Library and Information Sciences*. 3rd. ed. 1106–1123.

Mohrhardt, F.E. 1962. "IAALD International Library". *Quarterly Bulletin of the International Association of Agricultural Librarians and Documentalists*, 7(2/3), 132–135.

Mohrhardt, F.E. 1967. *Meeting National and International Library Needs*. [Schwing Library Lecture Series, No. 9]. Baton Rouge, LA: Louisiana State University.

O'Hara, Kieron and Wendy Hall. 2009. "Semantic Web". *Encyclopedia of Library and Information Sciences.* 3rd. ed. 4663–4676.

Poynder, Richard. 2002. "George Soros Gives $3 Million to New Open Access Initiative". *Information Today*, 19(4):20.

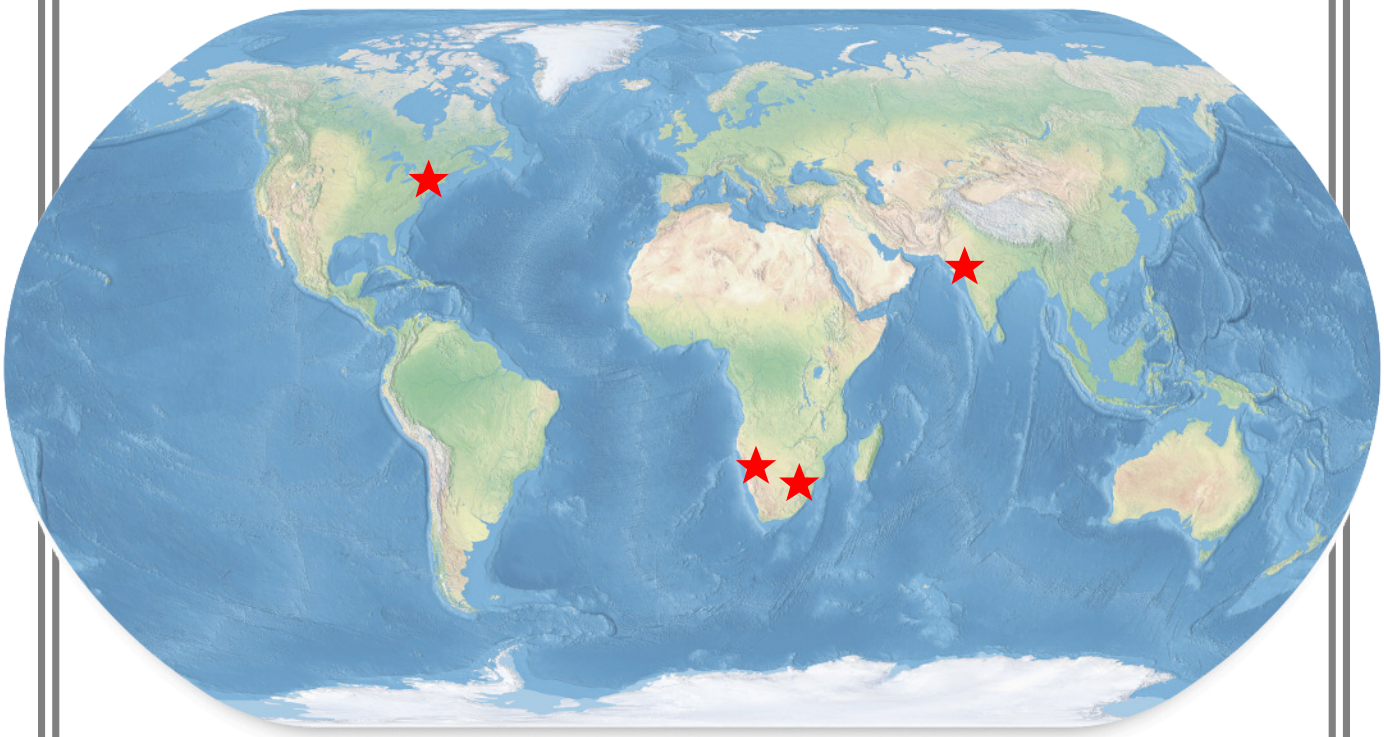
Schöpfel, Joachim and Dominic J. Farace. "Grey Literature". *Encyclopedia of Library and Information Sciences.* 3rd ed. 2029–2039.

IAALD is on the Move!

IAALD will be sponsoring events around the world in 2011, 2012 and 2013. We will be going from Africa to India to the USA with some exciting venues.

- This summer IAALD joins the Association for International Agricultural and Extension Education (AIAEE – *www.aiaee.org*), AGRISSON, and INFITA in hosting an international conference in Windhoek, Namibia.
- In the spring of 2012, IAALD Africa (*http://www.iaald-africa.org/*) will hold its biennial conference in Johannesburg and in the fall our colleagues in India will be hosting us for an IAALD international conference in Ahmedabad, India (*http://www.ahmedabad.org.uk/*).
- In 2013, Cornell University (*http://www.cornell.edu/*) in Ithaca, New York, USA will be hosting the XIVth IAALD World Congress.

Mark your calendar, save your money, and dust off your bags. ***We hope to see you at one of our events.***



IAALD Calendar

July 3-7, 2011 – ***Sustainable Value Chain: Agriculture for Food Security and Economic Development,*** an IAALD joint conference with AIAEE, AGRISSON and INFITA, Windhoek, Namibia.

May 23, 24, and 25th, 2012 – ***IAALD Africa Chapter Biennial meeting,*** Emperors Palace (tentative), Johannesburg, South Africa

October/November, 2012 – ***IAALD International Conference,*** Ahmedabad, India

July, 2013 – ***XIVth IAALD World Congress,*** Cornell University, Ithaca, New York, USA

Anton Mangstl: A Tribute

Antoinette Paris Greider

Author's Note: **Thank you to Robert Portegies-Zwart for providing much of the information in this article.**

At the end of March, the Food and Agricultural Organization of the United Nations (FAO) says goodbye to one of its leaders in information management, Dr. Anton Mangstl, Director of the Office of Knowledge Exchange, Research and Extension (OEK) and past IAALD Board member. Dr. Mangstl moved FAO's information function from one of a traditional library and indexing services to one of information management and document availability. He has moved information management beyond libraries with labor intensive processes to providing access to housed documents by taking FAO's documents directly to the user while reducing the labor required to provide the access and service.

Dr. Mangstl arrived at the FAO in 1996 to become Director of the Library and Documentation Systems Division (GIL) which was comprised of the David Lubin Memorial Library and the AGRIS/CARIS Information System. Dr. Mangstl soon went to work and in 1998 the World Agricultural Information Centre (WAICENT) was born. Although WAICENT had been in incubation for a number of years, it was under Dr. Mangstl's leadership that FAOSTAT, FAOINFO, and FAOSIS came together and became interactive and complementary. It was also a paradigm shift for FAO from paper to electronic distribution and distributed versus centralized storage of materials.

With this paradigm shift came a new name for the Division and GIL became the Knowledge Exchange and Capacity Building Division (KCE) with a new focus. Dr. Mangstl encouraged the organization to learn from, and give a voice to, the agricultural community regarding a wide range of issues affecting food security and agricultural production. Under his leadership the Division led the multi-agency Information Management Resource Kit e-learning initiative (IMARK) and the public/private partnership initiative on Access to Global Online Research in Agriculture (AGORA), a project that enables institutions in the poorest countries to access leading food and agriculture journals for little or no cost. Dr. Mangstl was responsible for overseeing the successful launch of FAO's strategic program on "Bridging the Rural Digital Divide" which lead to the development of the global e-Agriculture Community of Expertise (e-agriculture.org). He served on the United Nations Group on the Information Society, and on the Strategy Council for the Global Alliance for ICT and Development.

Anton Mangstl

In 2010, KCE became OEK — Office of Knowledge Exchange, Research and Extension furthering transforming FAO into a Knowledge Organization through the OEK services. The products and services include products for information and knowledge sharing, e-learning and knowledge networking, multi-lingual cross-media publishing, communication for development, copyright protection and preservation of FAO intellectual property. The Virtual Extension, Research and Communication Network strengthened linkages among rural institutions and farmers, while the Technologies for Agriculture (TECA) platform facilitated the exchange of proven agriculture technologies for small holders. His final initiative has been CIARD (Coherence in Information for Agricultural Research for Development).

OEK hosts the Secretariats of two bodies that have long-standing relations with FAO: the Global Forum on Agricultural Research (GFAR), and the Independent Science and Partnership Council, of the Consultative Group on International Agricultural Research (CGIAR), providing excellent opportunities to strengthen FAO's partnership with these organizations.

Additonal partnerships continue through IMARK and

Comparison of FAO Information Management in 1996 and 2011

Information Management in FAO — **1996**

Library and Documentation Systems Division (GIL)

- David Lubin Memorial Library
- AGRIS/CARIS Information System

Information Management in FAO — **2011**

Office of Knowledge Exchange, Research and Extension (OEK)

- **Branches**
  - Knowledge and Capacity for Development Branch (OEKC)
  - Knowledge Management and Library Services Branch (OEKM)
  - Publishing Policy and Support Branch (OEKP)
  - Research and Extension Branch (OEKR)
- **Hosted Secretariats**
  - Secretariat of the The Global Forum on Agricultural Research (GFAR)
  - Secretariat of the Independent Science and Partnership Council (ISPC) to the CGIAR

have facilitated adoption of e-learning techniques by FAO's technical program in food security, food safety, right to food, impact assessment, and plant genetic resources. The Office has partnered with other UN agencies, universities and training centers to include FAO's e-learning courses as part of certified training and university programs.

Dr. Mangstl's roots are in the academic community. He received his M.Sc. and PhD in Agronomy from Technische Universität in Munich, Germany and later served as Deputy Head to the Agronomy and Informatics Department of the same university. In 1989, he became the Director of ZADI (Centre for Agricultural Documentation and Information), in Bonn, Germany prior to becoming Director of GIL. In 2001, Dr. Mangstl was awarded Honorary Professor by the National Agriculture University of Ukraine.

This article chronicles a career of many achievements of an innovative information professional but I cannot leave with out saying something about Anton Mangstl as a person. I have had the privilege of knowing Dr. Mangstl personally since I began my tenure on the IAALD Board in 1990. While IAALD was still mired in the organization's past accomplishments, Dr. Mangstl could see what the organization needed to do. He was the "documentalist" part of IAALD that could see we needed to move away from passive delivery of information to active delivery of information. He believed in the profession and worked to make it better. He was one of the architects of the IAALD New Technologies Symposiums that were held in Washington, DC, and Bonn, Germany. It was here that many of us had a glimpse of what the future held in information management. In 2000, Dr. Mangstl made a trip to my library at the University of Kentucky and we spent a Friday visiting the libraries on campus and then a delightful Saturday exploring the rural areas of Kentucky. He was a wonderful guest and my then fiancé and I enjoyed our time with him.

It has been an honor for me to put together this article on Anton Mangstl. He has been an ardent supporter of IAALD and its mission. He has moved the information profession away from document storage to direct document delivery in an open environment. He has also been a good friend to both IAALD and me over the years. We wish him and his lovely wife Sophie well on this new adventure in their life.

Sources Consulted:

Mangstl, Anton. "A New Direction for FAO's Information Services World Agricultural Information Centre (WAICENT). http://www.cnshb.ru/aw/iaald_news_cee/5_3-97.htm [accessed 23 February 2011].

"Anton Mangstl". GFAR, Global Forum on Agricultural Research. http://www.egfar.org/egfarW/website/aboutgfar/management/person?contentId=3085 [accessed 23 February 2011].

FAO. "Office of Knowledge Exchange, Research and Education." http://www.fao.org/oek/oek-home/en/ [accessed February 23, 2011].

Selected Information Initiatives Developed During Anton Mangstl's Leadership

Access to Global Online Research in Agriculture (AGORA) – *http://www.aginternetwork.org/en/*

Bridging the Rural Digital Divide – *http://www.fao.org/rdd*

CIARD – *http://222.ciard.net*

e-agriculture – *http://www.e-agriculture.org*

Information Management Resource Kit (IMARK) – *http://www.imarkgroup.org/*

Research4life – *http://www.research4life.org*

Interested in Learning More about Repositories?
Check out the IMARK e-learning module described below in English, French and Spanish.

English:

The newest Information Management Resource Kit (IMARK) e-learning module entitled *Digital Libraries, Repositories and Documents*, offers learners the essentials for creating and managing their own digital libraries and repositories. It covers all the relevant processes for planning and resourcing, and considers the latest technologies and trends for digital data management and preservation. Relevant issues concerning the legal framework for copyright and intellectual property rights are also covered.

The module is offered online and on CD-ROM free of charge, and is currently available in English only. The Spanish version will be published in June 2011 and the French at a later date. For more information please visit the IMARK website: www.imarkgroup.org This module was developed by FAO, with support from the UK Department for International Development (DFID) and the International Network for the Availability of Scientific Publications (INASP), in the context of better "Coherence in Information for Agricultural Research for Development" (CIARD) (www.ciard.net), which is a unique global partnership in support of research communication in agriculture.

Française:

Le dernier module d'apprentissage électronique du Kit de ressources pour la gestion de l'information (IMARK) est intitulé : *Bibliothèques, archives et documents numériques*. Il expose les principes qui permettent de créer et de gérer des bibliothèques et référentiels numériques. Il explique tous les processus utiles à la planification et à la gestion des ressources, et présente les dernières technologies et tendances en matière de gestion et de préservation de données numériques. Les questions pertinentes concernant le cadre juridique du copyright et des droits d'auteurs sont également présentées.

Le module est disponible gratuitement en ligne et sur cédérom. Il est pour le moment disponible uniquement en anglais. La version espagnole sera publiée en juin 2011 et la version française à une date ultérieure. Pour plus d'informations, veuillez consulter le site web d'IMARK: www.imarkgroup.org

Ce module a été développé par la FAO, avec le soutien du Département du développement international du Royaume-Uni (DFID) et de International Network for the Availability of Scientific Publications (INASP), dans le cadre de l'initiative "Cohérence de l'information sur la recherche agricole pour le développement" (CIARD) (www.ciard.net), qui est un partenariat mondial unique à l'appui de la communication de la recherche dans l'agriculture.

Espanole:

El nuevo módulo de *e-learning* (aprendizaje electrónico) denominado *Bibliotecas, repositorios y documentos digitales,* brinda los elementos esenciales para aprender a crear repositorios y bibliotecas digitales. Dicho módulo forma parte del Repertorio de Recursos para la Gestión de la Información (IMARK, por sus siglas en inglés), el cual abarca todos los procesos relevantes para la planificación y asignación de recursos, considerando a la vez lo más reciente en tecnologías y tendencias para la gestión de los datos digitales y su conservación. Cubre además los aspectos más relevantes relacionados con el marco jurídico del derecho de autor y sobre la propiedad intelectual.

El módulo se ofrece en línea y en CD-ROM de forma gratuita y actualmente está disponible en idioma inglés. A partir de junio 2011, también se ofrecerá en español. Para obtener más información, visite el sitio web del IMARK: www.imarkgroup.org

Este módulo fue desarrollado por la Organización de las Naciones Unidas para la Agricultura y la Alimentación (FAO), con el apoyo del Departamento Británico para el Desarrollo Internacional (DFID), la Red Internacional para la Disponibilidad de Publicaciones Científicas (INASP), en el contexto de mejorar "la coherencia en la información para la investigación agraria para el desarrollo" (CIARD) (www.ciard.net), una asociación mundial única que apoya la comunicación de la ciencia y investigación en la agricultura.

AGRIS—From a Bibliographical Database to a Web Data Service on Agricultural Research Information

Angela Fogarolli, Dan Brickley, Stefano Anibaldi and Johannes Keizer

Editor's Note: **Paper presented at the IAALD XIIIth World Congress, Scientific and Technical Information and Rural Development, Montpellier, 26–29 April 2010.**

Abstract: **AGRIS provides a large collection of bibliographic references, such as research papers, studies and thesis, each including metadata such as conferences, researchers, publishers, institutions, and keywords from different thesauri such as AGROVOC. With the rise of the full text search and online availability of more research material, the role of bibliographic metadata appears redundant. When it is considered as a form of modeling that emphasizes relationships, connections and links, bibliographic metadata grows in value as the Web grows in connectivity. It can provide researchers with a map of the global research community, linking formal outputs (papers, data) with a wider grey literature (preprints, drafts) and with communication platforms (blogs, forums) that help researchers put formal findings into a wider context. This paper describes the evolving role of the AGRIS bibliographic database into becoming a hub of agricultural research literature. The large database of 3 million agricultural resources, collected by more than 150 institutions over the last 35 years, becomes the starting point to access the diverse knowledge in agricultural science and technology available globally on the web.**

Resumé: **AGRIS fournit une grande collection de références bibliographiques, comme les articles de recherche, études et thèses, qui comprennent des métadonnées comme les conférences, chercheurs, éditeurs, institutions, et mots-clés de différents thésaurus tel AGROVOC. Avec la hausse de la recherche sur plein texte et de la disponibilité en ligne de plus de matériel de recherche, le rôle des métadonnées bibliographiques apparaît redondant. Quand elles sont considérées comme une forme de modelage qui souligne des relations, les connexions et les liens, la valeur des métadonnées bibliographiques augmente avec l'augmentation de la connectivité du Web. Elles peuvent fournir aux chercheurs une carte de la communauté de recherche globale, reliant des productions formelles (articles et données) avec une littérature grise plus large (pré-imprimés, brouillons), avec des plateformes de communication (blogs, forums) qui aident les chercheurs à mettre des conclusions formelles dans un contexte plus large. Cet article décrit l'évolution du rôle de la base de données bibliographiques AGRIS, qui devient un moyeu de la littérature sur la recherche agricole. La grande base de données de 3 millions de ressources agricoles, alimentée par plus de 150 institutions durant ces 35 dernières années, devient le point de départ pour accéder à la diverse connaissance des sciences et technologies agricoles disponibles mondialement sur le Web.**

Resumen: **AGRIS ofrece una gran recopilación de referencias bibliográficas, tales como artículos de investigación, estudios y tesis. Cada una incluye metadatos como conferencias, investigadores, editores, instituciones y palabras clave de diferentes tesauros como AGROVOC. Con el aumento de búsquedas de texto completo y la disponibilidad en línea de un mayor volumen de material de investigación, la función de los metadatos bibliográficos parece redundante. Cuando se consideran como una forma de modelación que enfatiza las relaciones, conexiones y enlaces, los metadatos bibliográficos crecen en valor en la medida que la Web crece en conectividad. Pueden proporcionar a los investigadores un mapa de la comunidad global de investigación, vinculando los resultados formales (artículos, datos) con la literatura gris (preimpresos, proyectos) más amplia y con plataformas de comunicación (blogs, foros) que ayudan a los investigadores a posicionar los resultados formales en un contexto más amplio. Este trabajo describe la evolución de la función de la base de datos bibliográfica de AGRIS para convertirse en un núcleo de literatura pertinente a la investigación agrícola. Esta gran base de datos de 3 millones de recursos agrícolas, recopilada por más de 150 instituciones en los últimos 35 años, se está convirtiendo en el punto de partida para tener acceso a los diversos conocimientos en ciencia y tecnología agrícola disponibles a nivel mundial en la Web.**

Through exploring the evolving role of databases such as AGRIS, it has become clear that the connectivity patterns among the different types of information described in the database (researchers, topics, institutes, places) can be better reflected online through a more explicit representation both in Web metadata and in user-facing Web sites. The distributed nature of the world described by AGRIS naturally fits a "linked data" deployment model, in which AGRIS becomes more than a document discovery portal — it becomes an entry point and map of the entire research landscape around some topic or theme.

The Linked-data [1] techniques foster the link between resources through the Web. Such an approach requires an emphasis on sharing identifiers, names and descriptions of key real-world and abstract objects other than the bibliographic materials themselves: conferences, workshops, research centres, researchers, subject themes, and homepages.

None of this is news to the bibliographic professional; such concerns have been at the heart of metadata work for years. What is new is the presence of tools (standards, software) and community trends (open linked data, open archives, RSS/Atom syndication) that allow the full potential of such link-oriented metadata to be exploited.

An Entity Centric Approach to Data Aggregation

The idea of shifting the web from a large graph of documents to a large graph of data has become more popular

What is OKKAM – http://community.okkam.org ?

The OKKAM project will allow web users and professionals to share with whoever they want what they know about anything in very much the same way the Web allows the same people today to share their documents and get them in a click.

since Tim Berners-Lee proposed the idea of the semantic web. Scientists and practitioners have invested a lot of effort to realize this vision, often trying to adapt and reuse models and techniques originated from more traditional areas like databases and Artificial Intelligence (AI). There is an important difference between traditional knowledge-based systems, and the current work aiming at reaching semantic computing at the web scale. This is the notion of global interlinking of distributed pieces of knowledge.

At the base of such interlinking and the resulting semantic interoperability of fragments of data, is the notion of identity of and reference to entities. Systems that manage information about entities (objects/individuals/instances and so forth.) commonly issue identifiers for these entities, just in the way relational databases issue primary keys for records. If these identifiers are generated by the information system itself, several issues that hinder interoperability and integration arise. Among these are a proliferation of identifiers taking place (the same object is potentially issued with a new identifier in several information systems) and adding of identifiers cannot be achieved (for example one identifier can denote different entities in different information systems). In addition reference to entities across information systems is very complicated or impossible, because there are no means to know how an entity is identified in another system.

To overcome this lack of data-level integration, OKKAM [2] proposes a global, public infrastructure, called Entity Name System (ENS), which fosters the systematic creation and reuse of identifiers for entities in the global space of information and knowledge. This "a priori" approach enables systems to reference the entities which they describe with a globally unique identifier, and thus create pieces of information that are semantically pre-aligned around those entities. Semantic search engines or integration systems are then able to aggregate information from distributed systems around entities in a precise and correct way. We call this the "entity-centric approach" to semantic interoperability, and the resulting information/knowledge space the "web of entities".

The *OKKAMization* Process

The OKKAMization is the process necessary to include entities in existing information sources in the web of entities. The process involves the identification of entities inside existing repositories and the creation of unique identifier (OKKAM ID) for entities which are not already present in the ENS system. Creating an OKKAM ID for an object means getting a unique identifier which is an unambiguous way to refer to that object. An OKKAM ID is a well formed URI which enables the object to semantically connect to other global resources.

What is a URI ?

Internet space is inhabited by many points of content. A URI (Uniform Resource Identifier; pronounced YEW-AHR-EYE) is the way you identify any of those points of content, whether it be a page of text, a video or sound clip, a still or animated image, or a program. The most common form of URI is the Web page address, which is a particular form or subset of URI called a Uniform Resource Locator (URL). A URI typically describes:

- The mechanism used to access the resource
- The specific computer that the resource is housed in
- The specific name of the resource (a file name) on the computer

For example, this URI:

http://www.w3.org/Icons/WWW/w3c_main.gif

identifies a file that can be accessed using the Web protocol application, Hypertext Transfer Protocol, *("http://")* that is housed on a computer named *"www.w3.org"* (which can be mapped to a unique Internet address). In the computer's directory structure, the file is located at the pathname of *"/Icons/WWW/w3c_main.gif."* Character strings that identify File Transfer Protocol FTP addresses and e-mail addresses are also URIs (and, like the HTTP address, are also the specific subset of URI called a URL).

— Taken from SearchSOA.com Definitions
http://searchsoa.techtarget.com/definition/URI
[accessed February 12, 2011]

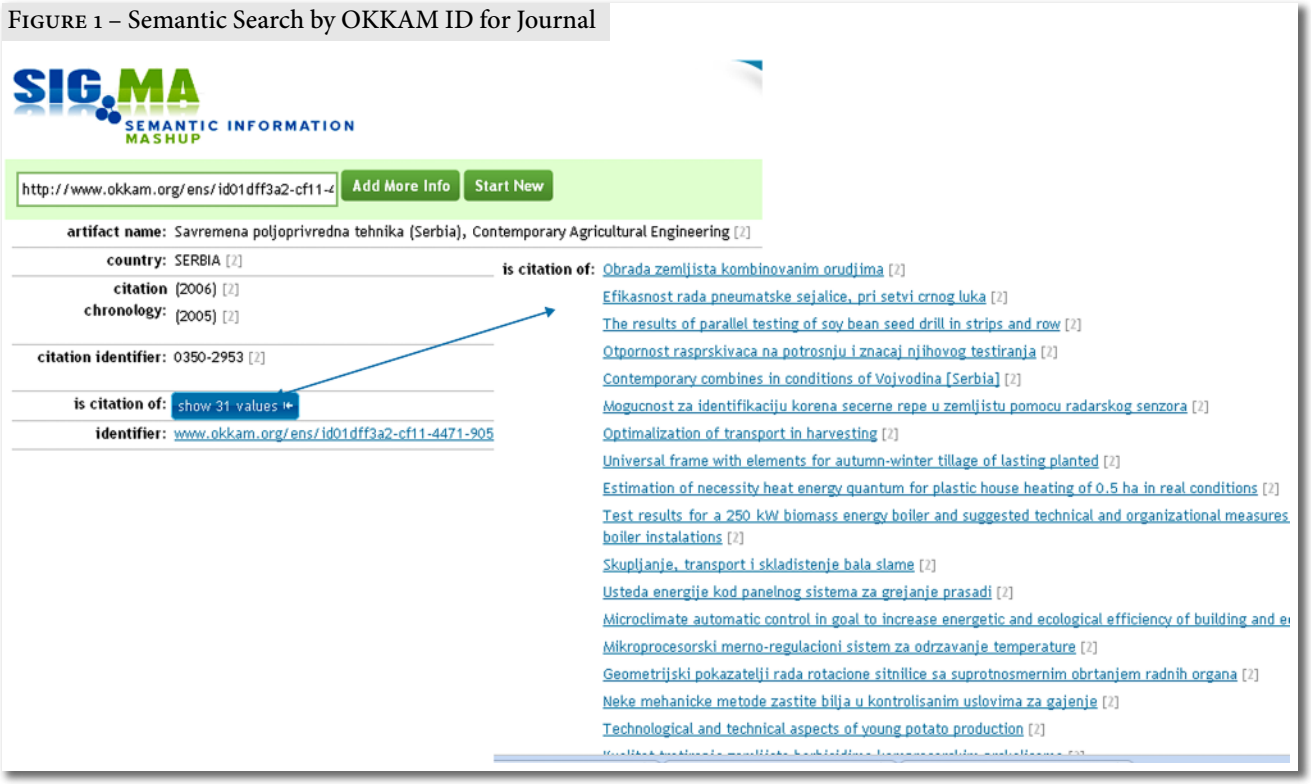
To allow for the correct creation of an OKKAM identifier it is necessary to collect a minimal set of information about an entity. This minimizes the risk of ambiguities. If, for example, to create an OKKAM ID about Mr. John Smith, by just using his name, the result will be an OKKAM ID that refers to a person. OKKAM will not be able to uniquely identify this entity because there are many "John Smiths" in the world. Building an OKKAM ID with more information such as state, city, or occupation allows for better recognition of the correct "John Smith."

The OKKAMization process of AGRIS repository is composed of four steps:

Step 1: Corpus entity recognition. This activity focuses on entity recognition inside the AGRIS repository and related sources.

Step 2: Associate OKKAM IDs to extracted entities. This task is based on the matching of extracted entities against the OKKAM ENS. If a match can be found for an entity then the identifier is re-used otherwise a new entity profile is created and thus a new unique identifier for the extracted entity.

Figure 1 – Semantic Search by OKKAM ID for Journal



Step 3: Enrichment of the AGRIS repository with OKKAM identifiers. The OKKAM identifiers generated in phase two are included in the XML files of the repository as another type of metadata. This allows for automatically identifying and aggregating entities inside the repository. The core point and objective of this phase is to enable entities based retrieval as well as to semantically connect entities in different contexts. From the user point of view this will translate into an efficient retrieval avoiding information overload.

Step 4: RDF enrichment of the AGRIS repository. This step has a big impact on the Web. It entails describing the AGRIS repository using the RDF notation. Publishing the repository using RDF makes the content of the repository understandable by external semantic search engines (SIG.MA, Google Project, and so forth). Therefore hidden semantic connections among entities can be discovered and showed to the user. Entities from the AGRIS repository can be described in RDF or micro formats in other web resources such as the FAO website and this will increase the semantic information that can be aggregated for the same entity.

What is RDF Notation?

The Resource Description Framework (RDF) is a family of World Wide Web Consortium (W3C) specifications originally designed as a metadata data model. It has come to be used as a general method for conceptual description or modeling of information that is implemented in web resources, using a variety of syntax formats.

— Taken from *Wikipedia* [accessed February 12, 2011]

The AGRIS Linked-data Model Using the OKKAM ENS Infrastructure

The AGRIS repository is a large and rich collection of bibliographic references encoded in a qualified DC XML format. Each XML document is structured in a metadata description for a resource which is sometimes available in PDF format.

In order to create a linked data model for AGRIS, the OKKAMization steps described in the previous section were followed.

Application of Step 1: Selected the journal entity type to illustrate the advantage of applying a linked data model to the AGRIS corpus beyond traditional identity types such as author.

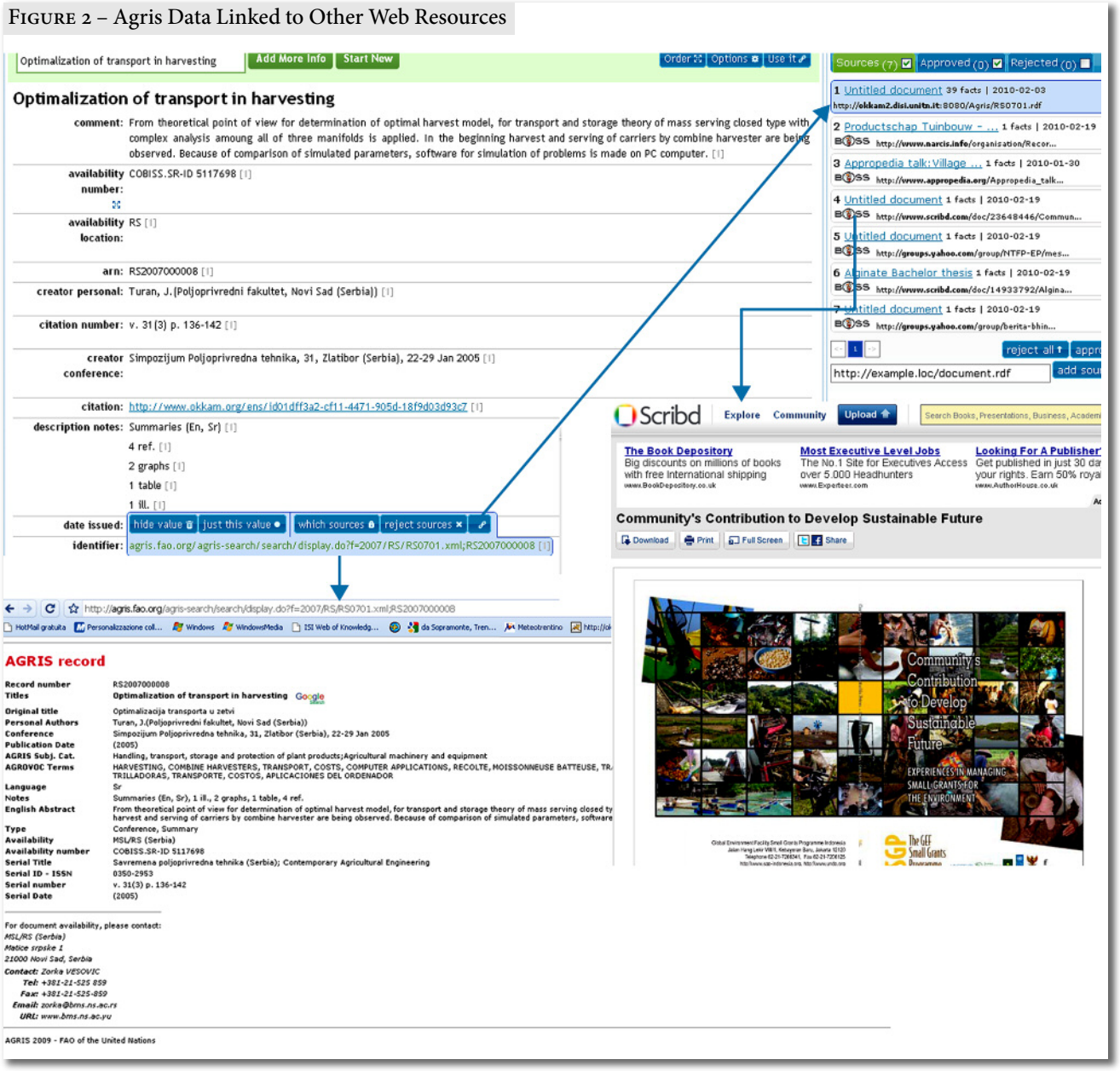
Application of Step 2: Obtained unique identifiers for each journal through OKKAM ENS.

Application of Step 3: Illustrated how unique identifiers are introduced in the original repository files.

Application of Step 4: Each file is translated into RDF format and submitted to a semantic web search engine.

The objective of assigning unique identifier to entities in the AGRIS repository leads to a light-weight data integration of entities and that enables inter-linkage among entities, which can come from different information sources as shown in Figures 1 and 2. As a result, efficient information retrieval is enabled within the AGRIS

FIGURE 2 – Agris Data Linked to Other Web Resources



repository and on the global scale by interlinking with other information sources.

Figures 1 and 2 show examples of the result of the OKKAMization process exposed in a Semantic Web Search engine.

Figure 1 shows the results of a search by unique identifier for a journal. The interface shows different statements about the journal resource and some of the attribute are clickable to enter in a deeper level of detail. In this example, a click on the "is citation" attribute displays all the article titles for that journal. A further click on one of the article titles displays the details of the article (Figure 2).

Information about the OKKAMized resource can be aggregated from different sources. The sources are displayed on the right side of Figure 2. A click on an attribute will give more details of the attribute itself. If the value of an attribute is a URL then this can connect with external information sources. As shown in Figure 2, it is possible from the article details to navigate to the AGRIS Website or to other external related sources. The amplitude of the inter-linkage with external resources grows with the use of the same unique identifier whenever that particular journal is mentioned in on the web.

As previously mentioned, the references of the AGRIS repository are encoded in a XML format. This type of file can be enriched with unique identifiers and this will allow the future representation of the unique identifier on the AGRIS web page enabling record linkage to the web of entities.

A snapshot of the XML of the AGRIS resource with a unique identifier for the journal in which the article appears is shown in Figure 3.

Following it in Figure 4 is an example of the automatic generated RDF file for each AGRIS resource article.

The unique identifier for the journal is presented in the colour red. The journal attributes are described inside the *rdf:Description* tag for the resource with a specific unique identifier (.rdf:about=”okkam_id value”).

Whenever the journal is cited inside the AGRIS record, the unique identifier will be used for its description. A reference to the original AGRIS website is presented in the colour green. This link will ensure reference to the AGRIS web search interface from any semantic web search engine.

Figure 3 – A snapshot of the XML of the AGRIS resource with a unique identifier for the journal in which the article appears

```
<ags:citation>
        <ags:citationTitle>Savremena poljoprivredna tehnika Serbia)</ags:citationTitle>
        <ags:citationTitle>Contemporary agricultural engineering</ags:citationTitle>
        <ags:citationIdentifier scheme="ags:ISSN">0350-2953</ags:citationIdentifier>
        <ags:citationNumber>v. 31(1-2) p. 29-37</ags:citationNumber>
        [1]<ags:citationChronology>(2005)</ags:citationChronology>
        <OkkamID>http://www.okkam.org/ens/ida53b7142-5880-4684-aab3-f83c2a6d0ea8</OkkamID>
</ags:citation>
```

Figure 4 – An example of the automatic generated RDF file for each AGRIS resource article

```
<?xml version="1.0"?>
<rdf:RDF
   xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
   xmlns:j.0="http://models.okkam.org/ENS-core-vocabulary.owl#"
   xmlns:j.1="http://purl.org/dc/terms/"
   xmlns:dc="http://purl.org/dc/elements/1.1/"
   xmlns:j.2="http://purl.org/agmes/1.1/" >
  <rdf:Description rdf:about="http://www.okkam.org/ens/id01dff3a2-cf11-4471-905d-18f9d03d93c7">
   <j.2:citationChronology>(2006)</j.2:citationChronology>
   <j.0:artifact_name>Savremena poljoprivredna tehnika (Serbia), Contemporary Agricultural
Engineering</j.0:artifact_name>
   <j.2:citationChronology>(2005)</j.2:citationChronology>
   <j.2:citationIdentifier>0350-2953</j.2:citationIdentifier>
   <j.0:country>SERBIA</j.0:country>
  </rdf:Description>
  <rdf:Description rdf:about="http://agris.fao.org/agris-
search/search/display.do?f=2007/RS/RS0701.xml;RS2007000023">
   <j.2:subjectThesaurus>PROPIEDADES TECNICAS</j.2:subjectThesaurus>
   <j.2:subjectThesaurus>TRITICUM</j.2:subjectThesaurus>
   <j.2:subjectThesaurus>http://www.fao.org/aos/agrovoc#c_2572</j.2:subjectThesaurus>
   <j.2:subjectThesaurus>http://www.fao.org/aos/agrovoc#c_2386</j.2:subjectThesaurus>
   <j.2:creatorConference>Simpozijum Poljoprivredna tehnika, 32, Zlatibor (Serbia), 28 Jan - 4 Feb
2006</j.2:creatorConference>
   <j.1:abstract>The paper shows presentation of the exploitational examination results for the wheat drill
sowing agregates. Some technical - technological drill solutions and the results of the working quality (norm,
drilling depth) and the exploitational parameters (working speed, output) have been shown.</j.1:abstract>
   <j.2:creatorPersonal>Mehandzic, R.(Poljoprivredni fakultet, Novi Sad (Serbia). Departman za poljoprivrednu
tehniku)</j.2:creatorPersonal>
   <j.2:creatorPersonal>Malinovic, N.(Poljoprivredni fakultet, Novi Sad (Serbia). Departman za poljoprivrednu
tehniku)</j.2:creatorPersonal>
   <j.2:subjectThesaurus>SEMIS EN LIGNE</j.2:subjectThesaurus>
   <dc:type>K</dc:type>
   <j.2:subjectThesaurus>SEMOIR</j.2:subjectThesaurus>
   <j.2:ARN>RS2007000023</j.2:ARN>
   <j.2:descriptionNotes>3 tables</j.2:descriptionNotes>
   <j.2:subjectThesaurus>TECHNICAL PROPERTIES</j.2:subjectThesaurus>
   <j.2:descriptionNotes>2 ref</j.2:descriptionNotes>
   <j.2:citation rdf:resource="http://www.okkam.org/ens/id01dff3a2-cf11-4471-905d-18f9d03d93c7"/>
```

Conclusions

In the AGRIS 2010 work [3], we have been prototyping a redesign for AGRIS that brings these concerns to the core of the system: both in our data modelling, and in the web presence. AGRIS will emphasize the networked, linked nature of the types of information it describes. AGRIS has for many years provided a large database of bibliographic references, such as research papers and theses, each including metadata such as conferences, researchers, institutions, and keywords using different thesauri such as AGROVOC.

For these reasons, the OKKAMization experiment that was explained above offers an effective and innovative solution for the global knowledge diffusion through semantic web technology. The solutions presented here allow the information retrieval system to perform stronger automatic elaboration offering data identification and aggregation. OKKAM allows the AGRIS repository to acquire the added value of making its full content available to the global web and at the same time to combine and aggregate information among and outside the organizational boundaries.

Notes and References

[1] Bizer C., T. Heath and T. Berners Lee, 2009. "Linked Data - The Story". ***International Journal on Semantic Web and Information Systems (IJSWIS),*** **5(3):1-22.**

[2] Bouquet, Paolo et al. 2008. "The back-bone of an open and scalable web of data" http://www.okkam.org/publications/stoermer-EntityNameSystem.pdf [accessed February 12, 2011]. Also published in the ***Second IEEE International Conference on Semantic Computing,*** **2008.**

[3] Brickley, D. et al. 2009. Designing AGRIS 2010 — ***Information linking and Agricultural Research.*** **ftp://fao.org/docrep/fao/012/ak689e/ak689e00.pdf [accessed February 12, 2011].**

Contact Information

Angela Fogarolli
OKKAM Project
University of Trento
Italy
afogarol@disi.unitn.it

Dan Brickley
Food and Agricultural Organization of the United Nations
Rome
Italy
dan.brickley@fao.org

Stefano Anibaldi
Food and Agricultural Organization of the United Nations
Rome
Italy
stefano.anibaldi@fao.org

Johannes Keizer
Food and Agricultural Organization of the United Nations
Rome
Italy
johannes.keizer@fao.org

DataStaR: A Data Sharing and Publication Infrastructure to Support Research

Gail Steinhart

Editor's Note: Paper presented at the IAALD XIIIth World Congress, Scientific and Technical Information and Rural Development, Montpellier, 26–29 April 2010.

Abstract: DataStaR, a Data Staging Repository (http://datastar.mannlib.cornell.edu/) in development at Cornell University's Albert R. Mann Library (Ithaca, New York USA), is intended to support collaboration and data sharing among researchers during the research process, and to promote publishing or archiving data and high-quality metadata to discipline-specific data centers and/or institutional repositories. Researchers may store and share data with selected colleagues, select a repository for data publication, create high quality metadata in the formats required by external repositories and Cornell's institutional repository, and obtain help from data librarians with any of these tasks. To facilitate cross-domain interoperability and flexibility in metadata management, we employ semantic web technologies as part of DataStaR's metadata infrastructure. This paper describes the overall design of the system, the work to date with Cornell researchers and their data sets, and possibilities for extending DataStaR for use in international agriculture research.

Resumé: DataStaR, un dépôt de données en développement (http://datastar.mannlib.cornell.edu) à la bibliothèque Albert R. Mann de l'Université de Cornell (Ithaca, New York, USA), a l'intention de soutenir la collaboration et le partage des données entre chercheurs pendant le processus de recherche, et de promouvoir la publication ou l'archivage de données et de métadonnées de haute qualité dans des centres de données avec des disciplines spécifiques, et/ou des dépôts institutionnels. Les chercheurs peuvent emmagasiner et partager des données avec les collègues choisis, choisir un dépôt pour la publication des données, créer des métadonnées de qualité supérieure dans les formats exigés par les dépôts externes et le dépôt de institutionnel de Cornell, et obtenir l'aide des bibliothécaires spécialistes des données pour n'importe laquelle de ces tâches. Pour faciliter l'inter-opérabilité entre les domaines et la flexibilité dans la gestion des métadonnées, nous employons des technologies sémantiques Web faisant parties de l'infrastructure des métadonnées de DataStaR. Cet article décrit la conception générale du système, le travail effectué jusqu'à ce jour avec les chercheurs de Cornell et leurs séries de données, et les possibilités pour élargir DataStaR à la recherche agricole internationale.

Resumen : DataStaR, un Repositorio de Clasificación de Datos (http://datastar.mannlib.corn ell.edu/) que se está construyendo en la Biblioteca Albert R. Mann de la Universidad de Cornell (Ithaca, Nueva York, Estados Unidos), busca apoyar la colaboración y el intercambio de datos entre investigadores durante el proceso investigativo y promover la publicación o el archivo de datos y metadatos de alta calidad en centros de datos o repositorios institucionales que sean específicos a cada disciplina. Los investigadores pueden almacenar y compartir datos con colegas seleccionados, seleccionar un repositorio para la publicación de datos, crear metadatos de alta calidad en los formatos requeridos por repositorios externos y el repositorio institucional de Cornell y obtener ayuda de los bibliotecarios especializados en datos con cualquiera de estas tareas. Para facilitar la interoperabilidad entre dominios y la flexibilidad en el manejo de los metadatos, se emplean tecnologías semánticas para la Web como parte de la infraestructura de metadatos de DataStaR. Este trabajo describe el diseño general del sistema, el trabajo que se ha realizado hasta la fecha con los investigadores de Cornell y sus conjuntos de datos, y las posibilidades para ampliar DataStaR para usarlo en la investigación agrícola internacional.

Sharing research data facilitates collaborative research among colleagues, and, when shared more widely, holds the potential to advance progress within a given discipline and even across disciplines. Research data may be used to reproduce and verify past results, plan future experiments, and support comparative studies and meta-analyses.

In spite of the potential benefits of sharing data, barriers to sharing exist. These may be cultural or sociological (sharing may not be the norm in certain disciplines; individuals may fear being "scooped"), procedural (confidentiality or commercialization concerns may mitigate against sharing), technological (suitable and accessible infrastructure may simply not exist), or logistical (researchers lack the skills and/or time to share their data). Our intention with the DataStaR — short for "Data Staging Repository" — project (http://datastar.mannlib.cornell.edu/) is to support the research process in a way that encourages data sharing more widely, primarily by reducing the significance of the last two barriers.

DataStaR: What is a Data Staging Repository?

DataStaR is both a platform and a set of services meant to facilitate data sharing in a way that is controlled by the researcher, as well as publication of data and metadata to appropriate repositories. We focus primarily on support for so-called "small science" data sets, those that don't require specialized infrastructure for storage, management, and access. DataStaR itself is only a temporary repository for data — working versions to be shared by colleagues, or final versions in preparation for submission to a permanent data repository. The notion of an intermediate working repository has other precedents. An infrastructure to support a curation continuum for research data, consisting of private, collaboration, and publication domains has been described and developed at Monash University in Australia by Treloar et al. (2007). Green and Gutmann (2007), in a paper describing the possibilities for partnerships between institutional repositories

and domain-specific repositories to encourage the migration of data from local to more widely shared environments, also get at this notion of a continuum or progression, and the types of support required to move works in progress to published versions. We emphasize that DataStaR is not a preservation repository, but is managed with long-term preservation of research data in mind (Steinhart *et al.*, 2009).

There are multiple benefits to the staging repository model. For users DataStaR offers a managed and controlled environment for collaboration with selected colleagues, off-site back up of valuable research data, tools to create metadata in a variety of formats, the ability to reuse information from previously created metadata, and assistance from librarians in determining an appropriate publication strategy and preparing data and metadata for publication. For librarians concerned with promoting responsible custodianship of research data created at their institution, the arrival of a new data set in DataStaR signals a curatorial opportunity. We see this combination as a potentially successful way to support the research process while simultaneously encouraging and supporting the publication of data sets to permanent repositories.

The DataStaR system consists of a Fedora-based repository (http://fedoracommons.org/) for storage of data set files; a semantic metadata store based on the Vitro software (http://vitro.mannlib.cornell.edu/) — a web-based ontology and instance editor developed at Mann Library; additional open-source components (DROID for file format identification, http://sourceforge.net/projects/droid/ and SWORD for deposit to some repositories, http://www.swordapp.org/); as well as custom code written specifically for this project (Figure 1).

A user may interact with the DataStaR system in the following ways: a researcher may upload a data set to the DataStaR repository, create minimal metadata, and assign permissions to grant access to data and metadata to selected colleagues, or the general public. At the time of upload, the user must indicate a destination repository for publication, although "to be determined" is a valid selection in the event a user is undecided or intends to use DataStaR solely for sharing data and not for publication. If no repository is selected, the user is presented with a simple (and optional) form for additional metadata. The selection of a specific destination repository triggers the display of a metadata form appropriate to that repository, although completion of this form is not required until the user is ready to publish the data set. Prior to or at the time of publication, the user completes the required metadata, consulting with project librarians as needed. The specifics of how a data set moves from DataStaR to a destination repository are varied, depending on the submission mechanism for that external repository. In some cases we are able to support direct deposit from DataStaR; in others, human mediation is necessary.

To facilitate cross-domain interoperability and flexibility in metadata management, we employ semantic web technologies as part of DataStaR's metadata infrastructure. Briefly, existing metadata schemas are converted to OWL ontologies and incorporated into the DataStaR system. [*Editor's note: For a description of OWL go to: http://en.wikipedia.org/wiki/Web_Ontology_Language.*] An advantage for users is that treating metadata as a collection of statements rather than static and stand-alone documents facilitates the reuse of previously

Figure 1 – Overview of DataStaR's system architecture, from left — initial access to the system by a data owner, to right — publication and distribution of data to other users. An access layer controls who may access the system and gives users the ability to grant access to others for their content. Users enter metadata about themselves and their research group as well as metadata for their data sets. The format of uploaded data files is determined by DROID and stored, along with other file-specific information, in the semantic metadata store, while data files are stored in a Fedora repository. For publication and direct distribution to users, XML metadata is written from the semantic metadata store. Data and metadata are downloaded or transmitted to users or archival repositories directly, or, in the case of some repositories, via the SWORD protocol.

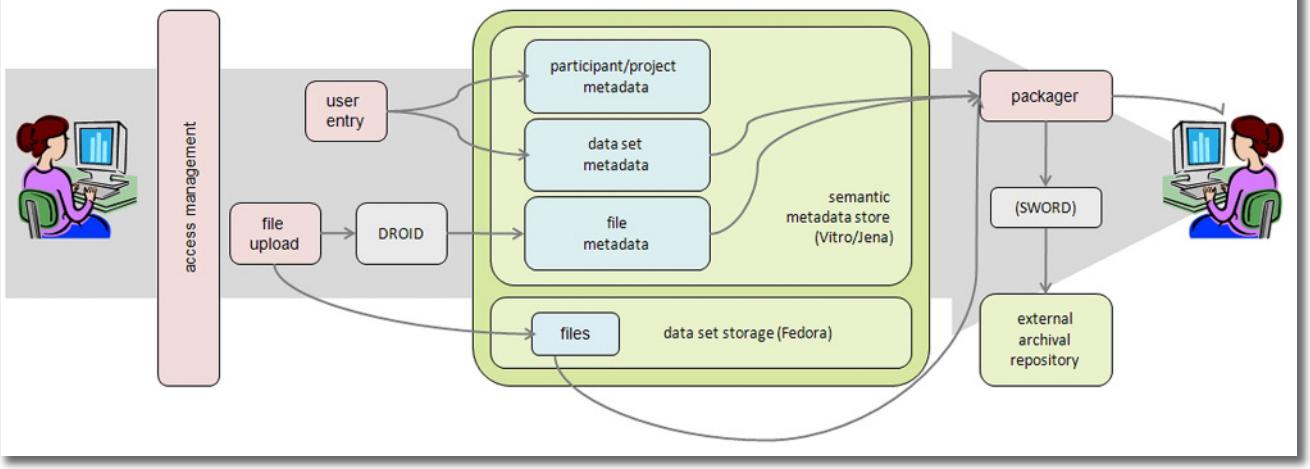


Table 1 – The repositories/domains and their metadata requirements supported in DataStaR's first round of development.

Repository or domain	Metadata requirements
eCommons (Cornell's institutional repository, http://ecommons.cornell.edu/)	DSpace/Dublin Core
Cornell University Geospatial Information Repository (CUGIR, http://cugir.mannlib.cornell.edu/)	Content Standard for Digital Geospatial Metadata (FGDC-CSDGM, http://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata)
Knowledge Network for Biocomplexity (KNB, http://knb.ecoinformatics.org/)	Ecological Metadata Language (EML, http://knb.ecoinformatics.org/software/eml/)
Virtual Center for Language Acquisition (VCLA, http://vcla.clal.cornell.edu/)	Open Language Archives Community (OLAC, http://www.language-archives.org/)

created statements in new metadata. It's not at all uncommon for a researcher to use the same field or laboratory methods, for example, or to conduct multiple studies in the same geographic location. Once that information has been entered in DataStaR to describe one data set, it is easily reused in the description of others. More broadly, we aim to support linked data in the future, and operate on the assumption that increasingly the application of semantic web approaches and technologies to the management of metadata will become standard practice. Our motivation for and approach to implementing a semantic approach to metadata management is described more fully in Lowe (2009).

In the first phase of development, currently underway, it's our goal to support publication to the repositories listed in Table 1. These repositories and standards reflect the needs of the researchers we work with; support for additional repositories and metadata standards may be added later, and will reflect the demands of the researchers with whom we work.

Providing Data Curation Services to Cornell Researchers

Currently, the DataStaR team is working with a number of research groups and individual researchers. These include Cornell's Upper Susquehanna River Basin Agricultural Ecology Program, the Cornell Biological Field Station, the Cornell Plantations Natural Areas Program, the Cayuga Lake Watershed Network, the Loon Project, and the Virtual Center for Language Acquisition. In addition, we plan to use DataStaR as a submission mechanism for data sets contributed to the Cornell University Geospatial Information Repository (CUGIR, http://cugir.mannlib.cornell.edu/).

The researchers involved with DataStaR are already motivated to share their data. Their motivations vary; in the case of the Upper Susquehanna River Basin Agricultural Ecology Program and the Cayuga Lake Watershed Network, sharing is motivated in part by a desire to make scientific findings available to managers, policy makers, and the general public. The Loon Project, as a recipient of funding from the US National Science Foundation's Long Term Research in Environmental Biology program, is explicitly required to disseminate its data, and is making use of DataStaR to do so. Long-term research also motivates sharing for the Cornell Biological Field Station, the Cayuga Lake Watershed Network, and the Cornell Plantations Natural Areas Program. Here, sharing previously collected data enables new research — whether by facilitating analysis over time, or simply providing background information to guide new research efforts. Facilitation of collaborative work within a research group is also a motivator: well-documented data, centrally accessible, makes it easier for collaborators to reuse and integrate data collected by others into their own research. In the case of the Virtual Center for Language Acquisition, this allows collaborative analysis of audio recordings, while in the case of the Upper Susquehanna River Basin Agricultural Ecology Program, sharing allows simulation modelers to validate their models using field-collected data. Because the groups we work with are already predisposed to share data, there is no resistance to sharing.

While collaborating researchers are motivated to share, they do have questions or concerns about the process, and most appreciate some level of assistance. The most commonly needed forms of assistance include help in deciding which data to share, help with data organization and formatting, and help with metadata creation. Deciding which data to share and how it should be organized depend to some extent on anticipated uses. For researchers collecting environmental data, anticipated reuse usually means analyzing data over time, or combining data sets from multiple researchers to perform comparative analyses. In these cases, data sets where the data have been somewhat processed are usually the most useful, rather than the raw data themselves, although raw data may allow others to check a researcher's intermediate calculations and final results. Decisions about organizing

data usually involve trade-offs that affect ease of use for the end user and ease of preparing and updating the data for the data owner. File format decisions also sometimes involve tradeoffs. Current and common proprietary formats may be easy to create and use, but are not suitable for long-term preservation, and may be incompatible with software other than that with which they were created. Non-proprietary formats, such as tab- or comma-delimited text files for tabular data, while they may not be in the working format that a researcher is accustomed to, are more stable in terms of long-term preservation and have greater potential for cross-platform compatibility now and in the future.

In terms of metadata, some metadata elements are fairly easy to understand and complete. Others may require specialized knowledge or an eye for details that researchers might reasonably overlook. Some examples of areas where expertise was provided to researchers in completing metadata include the use of controlled vocabularies for keywords and subject terms, assistance with crafting language for intellectual rights statements, and adherence to established conventions for specifying geographic coordinates.

Prospects for Applications in International Agriculture

We recognize the existence of well-established systems such as the International Crop Information System (ICIS) (Fox and Skovmand , 1996), and its crop-specific instances such as the rice (IRIS) and wheat (IWIS) systems (e.g. McLaren *et al,*. 2005). Systems like ICIS have certain advantages for agronomic data, including some degree of standardization of data that facilitates interoperability among data sets, and tools and applications for data use and analysis — capabilities that DataStaR lacks because it was developed to manage much more heterogeneous data sets.

Nevertheless, the following elements of DataStaR might be useful in some research contexts:

- infrastructure for preliminary and controlled data sharing during the research process,
- tools for documenting and moving data from that preliminary (staging) environment into the publication domain, and
- semantic approaches to metadata management.

Understanding its applicability requires consideration of one or more questions in each of these areas.

While we're not deeply familiar with the norms and practices for research in this area, we speculate that a data staging repository might work well to facilitate collaboration among researchers, particularly if they are distributed geographically, and if internet connectivity is somewhat reliable. A shared data repository can make it easier for researchers to ensure they are working from the same version of a data set, and also serve as a remote back-up.

There are two main requirements for the documenting and publishing function of DataStaR to be useful in this arena. The first is the existence of suitable destination repositories for agricultural research data. For researchers already participating in efforts such as ICIS or other systems, there may be no particular advantage to adding an intermediate layer to the process of publishing data, above and beyond the opportunities afforded by sharing works in progress with selected colleagues in the pre-publication stage. We don't know whether there are other repositories, institutional or discipline-specific, that might be usefully linked to a staging repository. If none exist, a single repository could serve both the staging and publication functions. The second requirement has to do with assistance to researchers in preparing and documenting data sets for publication. The reason this intermediate infrastructure works well at Cornell is that staff are available to guide researchers through the process of documenting and preparing data for publication, and the staging repository works well as a shared workspace where this preparatory work can be accomplished. Support for researchers using the system is as important a part of the system as the technological infrastructure itself.

Finally, a semantic approach to metadata management is, technologically, one of the most innovative aspects of DataStaR. The immediate benefit to users is the ability to reuse their own information, rather than re-entering or copying statements from one metadata record to another. Additional benefits of managing metadata in this way are more likely to be realized in the future, when linked data and infrastructure to support it are more common. Early adopters, while perhaps incurring some additional overhead in supporting a somewhat more complex infrastructure (compared to implementing an existing out-of-the-box repository solution) will avoid future costs of retrospective conversion of metadata to support the semantic web. The immediate utility of this approach depends on whether a flexible data sharing and publication environment capable of supporting multiple standards has value for the community, or whether useful links to other agricultural information systems already employing semantic web technologies can be made.

Conclusion

While we have worked with only a handful of research groups, we're pleased with their response to our services, which seem to fill some very real needs at Cornell. Since the project began, several researchers have asked us for assistance with data archival and dissemination plans in grant proposals. Furthermore, the University of Melbourne is in the process of adapting the DataStaR software for use as a data registry for the Australian National Data Service, having already implemented the core Vitro software as an expertise directory at the University of Melbourne.

Acknowledgements

The author gratefully acknowledges the contributions of the DataStaR team: Brian Caruso, Kathy Chiang, Jon Corson-Rikert, Dianne Dietrich, Ann Green, Brian Lowe, and Janet McCue, as well as Mary Ochs for providing comments on a draft of this paper.

This material is based upon work supported by the National Science Foundation under Grant No. III-0712989. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Fox, P. N., and B. Skovmand. 1996. The International Crop Information System (ICIS) — Connects Genebank to Breeder to Farmer's Field. In: Plant adaptation and crop improvement. M. Hammer and Cooper G.L., eds. Wallingford, UK: CAB International. p.317–326.

Green, A.G., and M.P. Gutmann. 2007. Building Partnerships Among Social Science Researchers, Institution-Based Repositories and Domain Specific Data Archives. OCLC Systems & Services 23(1): 35–53.

Lowe, B. 2009. DataStaR: Bridging XML and OWL in Science Metadata Management. Metadata and Semantic Research 46: 141–150.

Mclaren, C. G., R. M. Bruskiewich, and A. M. Portugal, A. B. Cosico. 2005. The International Rice Information System. A Platform for Meta-analysis of Rice Crop Data. Plant Physiology 139(2): 637–42.

Steinhart, G., D. Dietrich, and A. Green. 2009. Establishing trust in a chain of preservation: The TRAC Checklist Applied to a Data Staging Repository (DataStaR). D-Lib Magazine 15(9/10).

Treloar, A., D. Groenewegen, and C. Harboe-Ree. 2007. The Data Curation Continuum: Managing Data Objects in Institutional Repositories. D-Lib Magazine 13(9/10).

Contact Information:

Gail Steinhart
Albert R. Mann Library
Cornell University, USA
GSS1@cornell.edu

Trends of the Institutional Repositories on Agricultural Universities in Japan

Takashi Nagatsuka and Naohisa Koremura

Editor's Note: **Paper presented at the IAALD XIIIth World Congress, Scientific and Technical Information and Rural Development, Montpellier, 26–29 April 2010.**

Abstract: **This paper discusses the present status of institutional repositories in Agricultural Universities in Japan as found in a survey conducted in January 2010. There are over seventy of agricultural universities in Japan which include the broad areas related to agriculture such as the faculty and graduate schools of Agriculture, Veterinary Medicine, Life Science, Fisheries Sciences, Agricultural Resource Sciences, Horticulture, Marine Science and Technology, Textile Science and Technology, and Environmental Studies. The experimental project of institutional repositories was started in 2004 and since then, over 100 universities have joined the National Institute of Informatics Institutional Repositories Program. The contents of institutional repositories consist of journal articles, dissertations, bulletins, meeting articles, documents for meetings, books, technical reports, magazine articles, preprints, learning materials, data/datasets, software and other materials. The number and type of contents of institutional repositories differ between each agricultural university. The future direction of institutional repositories of agricultural universities in Japan is also discussed and concludes the paper.**

Resumé: **Cet article discute le statut présent des dépôts institutionnels dans les universités agricoles au Japon, selon une enquête dirigée au mois de janvier 2010. Il y a plus de soixante-dix universités agricoles au Japon qui couvrent les larges secteurs liés à l'agriculture comme la faculté et les universités en agriculture, médecine vétérinaire, science de la vie, sciences des pêches, sciences des ressources agricoles, horticulture, science et technologie marine, science et technologie du textile, et études environnementales. Le projet expérimental de dépôts institutionnels a commencé en 2004 et depuis, plus de 100 universités ont joint le Programme de dépôts institutionnels de l'Institut national d'informatique. Les contenus des dépôts institutionnels contiennent des articles de journaux, des thèses, des bulletins, des articles et documents de réunions, des livres, des rapports techniques, des articles de revue, des pré-imprimées, du matériel de formation, des données/sets de données, des logiciels et autres matériels. Le nombre et le type de contenus des dépôts institutionnels diffèrent entre chaque université agricole. La direction future des dépôts institutionnels des universités agricoles au Japon est aussi discutée et conclut l'article.**

Resumen: **Este trabajo trata la condición actual de los repositorios institucionales de las universidades agrícolas del Japón, con base en los resultados de una encuesta realizada en enero del 2010. Hay más de 70 universidades agrícolas en Japón, cuyas facultades abarcan áreas amplias relacionadas con la agricultura tales como Agricultura, Medicina Veterinaria, Ciencias Biológicas, Ciencias de la Pesca, Ciencias de los Recursos Agrícolas, Horticultura, Ciencia y Tecnología Marina, Ciencia y Tecnología Textil, y Estudios Ambientales. El proyecto experimental de repositorios institucionales comenzó en el 2004 y, desde entonces, más de 100 universidades se han unido al Programa de Repositorios Institucionales del Instituto Nacional de Informática. El contenido de los repositorios institucionales consta de artículos de revista, tesis de grado, boletines, actas de reuniones, documentos para reuniones, libros, informes técnicos, pre-impresos, materiales didácticos, datos/conjuntos de datos, software y otros materiales. El número y tipo de contenido en los repositorios institucionales difieren para cada universidad agrícola. También se analiza la dirección futura de los repositorios institucionales de las universidades agrícolas del Japón y se presentan las conclusiones.**

Institutional repositories provide universities with a sound method of significantly improving scholarship and scholarly communication. Japanese university-based institutional repository provides a service to universities in Japan for the management and dissemination of digital materials created by the institution and its community members. While operational responsibility for these services may reasonably be situated in different organizational units at different universities, an effective institutional repository, of necessity, represents a collaboration among librarians, information technologists, archives and materials managers, faculty, and university administrators and policymakers (Lynch, 2003).

An institutional repository differs from other digital collections that might be offered by a university library such as access to e-books or e-journals. In the case of an institutional repository, content is deposited in a repository by the creator or owner, so it is necessary that repository architecture manages the content and the metadata, and repository software offers a minimum set of basic services to input or retrieve the content. The repository must be sustainable, trusted, well supported and well-managed (Tedd, 2009). At any given point in time, an institutional repository will be supported by a set of information technology (IT) tools. A key part of the services that comprise an institutional repository is the management of technological changes, in particular the migration of digital content from one set of IT tools to the next as part of the organizational commitment to providing repository services. It can be said that an institutional repository is more than a fixed set of software and hardware (Lynch, 2003). For a scientist who wants to get the widest possible dissemination, and to gain as much recognition as possible, open access is the way to

go. The visibility comes from having open access vehicles indexed by Google, Google Scholar and other web search engines. The recognition and impact come because search engines lead to direct access to the scientist's full-text articles on the readers' desktops. There are two main avenues to make open access work. One is to publish articles in open access journals and the other is to deposit articles in open access research repositories (Swan, 2009). The development of institutional collections of published outputs which are made openly available over the Web is an important role of the university-based institutional repository. From the perspective of the institutions, the research institutional repository could have two functions. One is an administration tool for the institutions in relation to annual reports, research assessment exercises, and so forth and the other is a way of showcasing the research output of the institutions. Clearly, the present state of research institutional repositories is not yet in the final phase of innovation implementation called 'infusion', whereby the scholarly communication system as a whole will function at a higher level. However, working on three tracks to improve the functionality of the repositories for authors, institutions and users will make this final stage of innovation adoption achievable (Vernooy-Gerritsen et al, 2009).

The experimental project of Japanese institutional repositories began in 2004 (Murakami and Adachi, 2006). To enhance the current content services and develop the next generation of scholarly content infrastructure, the National Institute of Informatics in Japan (NII) helped Japanese universities to introduce and operate institutional repositories (NII Phase 1 Report, 2008). Over 100 Japanese Universities have joined the National Institute of Informatics Institutional Repositories Program (NII Institutional Repositories Program http://www.nii.ac.jp/irp/en/about/) since the beginning. There are over seventy Japanese agricultural faculties and universities in the broad areas related to agricultural studies (agriculture, veterinary medicine, life science, fisheries sciences, agricultural resource sciences, horticulture, marine science and technology, textile science and technology, environmental studies). The contents of these repositories include journal articles, dissertations, bulletins, meeting articles, meeting documents, books, technical reports, magazine articles, preprints, learning materials, data/datasets, software, and other materials.

This paper evaluates the present status of institutional repositories of agricultural faculties and universities in Japan. The study highlights five main factors influencing the status of institutional repositories, the type of university, the type of materials, the number of titles, the number of titles in relation to the size of university. The study looks at the difference in the number and type of contents among institution repositories in agricultural faculties and universities.

Methodology

A list of institutional repositories in Japan that included Japanese universities was obtained from the NII Institutional Repositories Program (http://www.nii.ac.jp/irp/en/about). In January of 2010, the authors accessed each university site based on the list and downloaded the metadata from each university. The datasets which included all of the elements from the list above were created from institutional repositories created by Japanese agricultural faculties and universities. In addition, the authors registered for a password to each site to enable them to download full-text articles or original documents, if needed, to verify the type of material.

Data Analysis and Results

Institutional Repositories at Universities in Japan – The total number of institutional repositories at Japanese universities and the number of Japanese universities who have established institutional repositories are shown in Table 1. The total number of universities is 773. This number is based on the data of the School Basic Survey of 2009 by the Ministry of Education, Culture, Sports, Science and Technology in Japan. The total number of repositories is 104 which is 13.5% of the institutions surveyed. The number of agricultural universities was 77 which is 10% of all universities. The number of institutional repositories established since 2004 was 39 or 50.6% of the universities in agricultural fields had established institutional repositories since 2004 (5% of the total of all universities). It is significant that more than half of the agricultural universities have established their institutional repositories in the last six years.

There are three different types of Japanese universities represented in the above table. They are National University Corporations, Prefectural and Municipal Universities, and Private Universities. The breakdown of repositories by type of institution is 73 for National University Corporation (70.2%), 27 for Private Universities (26.0%), 4 for Prefectural and Municipal Universities (3.8%) respectively. The number and percentage of institutional repositories at faculties and universities in agricultural fields among the three different types of university were 37 (94.9%) for the National University Corporation, 1 for Prefectural and Municipal Universities (2.6%) and 1(2.6%)

Table 1 – Number of institutional repositories in Japanese Universities

	Number	%
All Japanese Universities	773	100
Agricultural Universities and Faculties	77	10
Total Institutional Repositories	104	13.5
Agricultural Universities and Faculties Repositories	39	5

Table 2 – Number of institutional repositories at the three different types of universities in Japan

	National University Corporation	%	Prefectural and Municipal University	%	Private University	%	Total	%
All Japanese Universities	86	11.1	92	11.9	595	77.0	773	100.0
Agricultural Universities and Faculties	43	55.8	12	15.6	22	28.6	77	100.0
Total Institutional Repositories	73	70.2	4	3.8	27	26.0	104	100.0
Agricultural Universities and Faculties Repositories	37	94.9	1	2.6	1	2.6	39	100.0

for the Private Universities. The largest percentage of institutional repositories in the National University Corporations are located in the agricultural fields (Table 2).

Material titles and types of materials

The types of materials that are archived in institutional repositories consist of journal articles, theses or dissertations, departmental bulletin/papers, conference papers, presentations, books, technical reports, research papers, articles, preprints, learning materials, data or datasets, software, and other materials. The number of titles in each in each material type for the different fields of agricultural study was surveyed to see whether the number and types of content of the institutional repositories in each agricultural university varied. The fields of agricultural study were divided into four broad groups; agriculture, veterinary science, animal husbandry and fisheries science. The highest percentage of deposits in a repository among the four groups was in the field of agriculture. These deposits represented 69.8% of the total with 208 titles in the various material groups. Of this group, departmental bulletins represented the greatest number of publications with theses and dissertations second and journal articles third. Departmental bulletins were also the highest number of publications deposited in the repository in fisheries science with 41.7% of the total items deposited while journal articles dominated Veterinary Sciences and Animal Husbandry. A summary of materials deposited by discipline can be found in Table 3.

Table 3 – Number of materials deposited by type in each field of agricultural studies in Japan

	Agriculture		Veterinary science		Animal husbandry		Fisheries science		Total	
Types of materials	Number of material titles	%	Number of material titles	%	Number of material titles	%	Number of material titles	%	Number of material titles	%
Journal Article	25	12.0	9	34.6	9	32.1	6	16.7	49	16.4
Thesis or Dissertation	33	15.9	5	19.2	2	7.1	5	13.9	45	15.1
Departmental Bulletin Paper	95	45.7	5	19.2	5	17.9	15	41.7	120	40.3
Conference Paper	17	8.2	—	—	1	3.6	4	11.1	22	7.4
Presentation	—	—	—	—	—	—	—	—	—	—
Book	4	1.9	—	—	—	—	—	—	4	1.3
Technical Report	17	8.2	7	26.9	9	32.1	1	2.8	34	11.4
Research Paper	2	1.0	—	—	—	—	—	—	2	0.7
Article	—	—	—	—	1	3.6	—	—	1	0.3
Preprint	—	—	—	—	—	—	—	—	—	—
Learning Material	6	2.9	—	—	—	—	2	5.6	8	2.7
Data or Dataset	2	1.0	—	—	1	3.6	—	—	3	1.0
Software	—	—	—	—	—	—	—	—	—	—
Others	7	3.4	—	—	—	—	3	8.3	10	3.4
Total	208	100.0	26	100.0	28	100.0	36	100.0	298	100.0

The types of materials which are archived in institutional repositories were surveyed at the sites in all Japanese universities as well as Japanese agricultural faculties and universities. The number of deposits in all universities is derived from the data at the NII Institutional Repositories Program site in January of 2010. The materials deposited in the agricultural faculties and universities were obtained by accessing each institutional repositories at the individual faculties and universities in January of 2010 by using the registered password and downloading the materials. The total number of items in institutional repositories for all universities 798,115 and 31,507 in agricultural faculties and universities which is only 3.9% of the total items deposited nationally. The trend in all universities is to deposit department bulletin/papers (43% of all deposited) and then journal articles (25% of all deposited) with everything else comprising the rest of the deposits. The same holds true for Japanese agricultural faculties and universities but by a larger percentage with two-thirds of the deposits being departmental bulletin/papers, 11% being journal articles and then theses and dissertations third. The detailed breakdown can be found in Table 4.

Table 4 – The types of materials archived in institutional repositories in Japan

Types of materials	All Universities		Agricultural Faculties and Universities	
	Number of materials	%	Number of materials	%
Journal Article	199,603	25.0	3,505	11.1
Thesis or Dissertation	43,309	5.4	3,036	9.6
Departmental Bulletin Paper	343,542	43.0	20,858	66.2
Conference Paper	52,796	6.6	2,401	7.6
Presentation	2,206	0.3		
Book	16,526	2.1	23	0.1
Technical Report	4,181	0.5	1,336	4.2
Research Paper	13,205	1.7		
Article	26,993	3.4	106	0.3
Preprint	294	0.0		
Learning Material	4,138	0.5	112	0.4
Data or Dataset	607	0.1	125	0.4
Software	8	0.0		
Others	90,707	11.4	5	0.0
Total	798,115	100.0	31,507	100.0

Materials in Each Field of Agricultural Studies

The number of materials in each field of agricultural study was surveyed to see whether the number and type of materials in institutional repositories in each agricultural university varied (Table 5). Once again the field of agriculture represented the highest number of materials deposited with 74.9% of the deposits (23,586). Departmental bulletin/papers were again the highest percentage of deposits with conference papers second and theses or dissertations third in the number of deposits. Departmental bulletin/papers dominated the deposits in the other three areas (veterinary science, animal husbandry, fisheries) with variations in what material type was second and third in numbers. A breakdown by subject can be found in Table 5.

Six Main National Universities and Other Universities in Agricultural Studies

The number and types of material which are archived in institutional repositories in the six main Japanese national universities and other Japanese universities in agricultural studies are shown in Table 6. Materials from the six main national universities comprise 48.4% (15,467) of the total items deposited. Other universities contributed 16,504 items or 51.6% of the total. As in the other 28 universities departmental bulletin/papers dominated the number of deposits with 79.3% of the total. Unlike the other 28 universities, conference papers rather than theses or dissertations were a distance second with 10.6% of the total. Journal articles constituted the third highest type of material deposited. Although the percentage of departmental bulletin/papers was the highest in both six national universities and the 28 other universities, the percentage of departmental bulletin/papers in six main national universities was highest at nearly 80%. This high percentage shows that the coverage of materials in the institutional repositories in the six main national universities is limited in scope. As can be seen in the in the percentage of "theses or dissertation" and journal article in the other 28 universities, it is clear that the variation of types of material is relatively larger than that of six main national universities.

Conclusion

It is clear that the institutional repository is a powerful tool that can serve as an engine of change in our institutions of higher education, and more broadly for the scholarly enterprises that they support. If properly developed, it advances a number of goals, and addresses an impressive range of needs. This is an area where universities

Table 5 – Number of materials in each field of agricultural studies in Japan

Types of materials	Agriculture		Veterinary science		Animal husbandry		Fisheries science		Total	
	Number of materials	%	Number of materials	%	Number of materials	%	Number of materials	%	Number of materials	%
Journal Article	2,285	9.7	397	13.8	213	13.0	610	17.9	3,505	11.1
Thesis or Dissertation	2,349	10.0	384	13.3	69	4.2	234	6.9	3,036	9.6
Departmental Bulletin Paper	15,512	65.8	2,037	70.8	775	47.2	2,534	74.5	20,858	66.2
Conference Paper	2,398	10.2	—	—	—	—	3	0.1	2,401	7.6
Presentation	—	—	—	—	—	—	—	—	—	—
Book	23	0.1	—	—	—	—	—	—	23	0.1
Technical Report	798	3.4	60	2.1	478	29.1	—	—	1,336	4.2
Research Paper	—	—	—	—	—	—	—	—	—	—
Article	—	—	—	—	106	6.5	—	—	106	0.3
Preprint	—	—	—	—	—	—	—	—	—	—
Learning Material	91	0.4	—	—	—	—	21	0.6	112	0.4
Data or Dataset	125	0.5	—	—	—	—	—	—	125	0.4
Software	—	—	—	—	—	—	—	—	—	—
Others	5	0.0	—	—	—	—	—	—	5	0.0
Total	23,586	100.0	2,878	100.0	1,641	100.0	3,402	100.0	31,507	100.0

need to invest aggressively, but where they also need to proceed thoughtfully and carefully, with broad consultation and collaboration across the campus community and with a full understanding that if they succeed they will permanently change the landscape of scholarly communication (Lynch, 2003). It is absolutely essential to expand the coverage of contents in the Japanese institutional repositories in the future to support the activities of academic research in the universities. Every institution in higher education will need or want to run an institutional repository in the future. The institutional repository in the Japanese university is now becoming a reality. The number and types of material in the repositories differ for each agricultural university. It will be necessary to add to the content of institutional repositories in each Japanese agricultural university in Japan to improve the access to academic information in the future. For example, the amount of learning materials in Japanese institutional repositories is small and should be expanded for use inside and outside of the classroom.

The institutional repositories in the agricultural fields in National University Corporations form the largest portion of materials deposited. The institutional repositories at Prefectural and Municipal Universities, and Private Universities should expand the content both in quality and

Table 6 – Number of materials in six main national universities and the other universities

Types of materials	Six main national universities		The other universities	
	Number of materials	%	Number of materials	%
Journal Article	860	5.6	2,645	16.0
Thesis or Dissertation	349	2.3	2,687	16.3
Departmental Bulletin Paper	12,273	79.3	8,585	52.0
Conference Paper	1,690	10.9	711	4.3
Presentation	—	—	—	—
Book	—	—	23	0.1
Technical Report	223	1.4	1,145	6.9
Research Paper	—	—	—	—
Article	—	—	106	0.6
Preprint	—	—	—	—
Learning Material	70	0.5	42	0.3
Data or Dataset	—	—	125	0.8
Software	—	—	—	—
Others	2	0.0	435	2.6
Total	15,467	100.0	16,504	100.0

quantity of content in the future. It is clear that the number and type of material titles are different among the four broad groups of agricultural studies that are archived in institutional repositories of agricultural faculties and universities. The field of agriculture provides the majority of the contents archived in institutional repositories but the majority of deposits are for papers in departmental bulletins in all cases with over 60% of the total deposits as compared to 45% in other universities. For this reason the coverage of institutional repositories in the agricultural fields is very limited compared with all areas of research in Japanese universities. The number and type of materials archived in institutional repositories at the agricultural faculties and universities are different among the four broad groups in agricultural studies but the fact remains that the percentage of departmental bulletin/papers was the highest in both cases in six main national universities (almost 80%) and the other universities in agricultural studies in Japan. It is obvious that the institutional repositories in the agricultural research fields have the potential to introduce a broad range of material members to the university in the future.

References

Lynch, C.A. 2003. Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age. ARL, 226 (February 2003): 1–7 http://www.arl.org/resources/pubs/br/br226/br226ir.shtml.

New Horizons in Academic and Scholarly Communication: NII National Institutional Repositories Program Phase 1 Report. 2008. http://www.nii.ac.jp/irp/en/archive/pdf/CSI_Phase1_Report.pdf

Murakami Y. and J. Adachi . 2006. Institutional Repositories in Japan. In: Lecture Notes in Computer Science 4312, 540–549. https://www.nii.ac.jp/irp/en/event/pdf/ICADL_2006.pdf

Swan A., 2009. The communication of science and technology in a sustainable future. World Review of Science, Technology and Sustainable Development. 6(1), 1–10.

Tedd, L.A., 2009. Open access publishing and institutional repositories: an overview. In E-content management: challenges and strategies. 7th International CALIBER conference. 572–584. URI: http://hdl.handle.net/2160/3986

Vernooy-Gerritsen, M., G. Pronk and M.V.D. Graaf. 2009. Three Perspectives on the Evolving Infrastructure of Institutional Research Repositories in Europe. Ariadne. 59, 30-April-2009. http://www.ariadne.ac.uk/issue59/vernooy-gerritsen-et-al/

Contact Information

Takashi Nagatsuka
Dept. of Library, Archival and Information Studies
Tsurumi University, Japan
nagatsuka-t@tsurumi-u.ac.jp

Naohisa Koremura
Scientific Information Program
Tokyo University of Agriculture, Japan
koremura@nodai.ac.jp.

Study on the National Agricultural Digital Library Information Integrated Service in China

Wenfei Zhuo, Shuchun Pan and Zongfang Xie

Editor's Note: **Paper presented at the IAALD XIIIth World Congress, Scientific and Technical Information and Rural Development, Montpellier, 26–29 April 2010.**

Abstract: This paper describes the National Agricultural Digital Library Information Integrated Service (DLIIS), a user-oriented information service platform serving China's agricultural science and technology sector, that seamlessly integrates a large volume of distributed content and information resources and techniques with a mechanism to enhance the cooperation of participating organizations. The integrated platform, which is part of the development of the digital library, synthesizes multifunctions of library services and gives the users the opportunity to access information when and how they want. The paper emphasizes the National Agricultural DLIIS System design and information integrated service platform through a discussion of the:

- **establishment of a user oriented National Agricultural DLIIS System;**
- **design of a four layer network architecture (user interactive, system application, distributed resource integration and technical support layer) of DLIIS according to the digital library's layer theory; and**
- **design of the service platform including resource integration and services cooperation.**

Resumé: Cet article décrit le Service d'information intégrée de la bibliothèque numérique agricole nationale (DLIIS), une plateforme de services d'information orientée vers l'utilisateur, servant le secteur de la science et de la technologie agricole de la Chine, qui intègre de façon continue un grand volume de contenu et de ressources et techniques d'information, distribués grâce à un mécanisme qui améliore la coopération des organisations participantes. La plateforme intégrée, faisant partie du développement de la bibliothèque numérique, synthétise les fonctions multiples des services de bibliothèque et donne aux utilisateurs l'occasion d'accéder à des informations quand et comment ils les veulent. L'article souligne la conception du système agricole national de DLIIS et de la plateforme de services d'information intégrée par une discussion:

- **sur l'établissement d'un système agricole national orienté vers l'utilisateur comme le DLIIS;**
- **sur la conception d'une architecture de réseau à quatre niveaux (l'utilisateur interactif, l'application de système, l'intégration de ressource distribuée et les niveaux de soutien technique) de DLIIS selon la théorie de la bibliothèque numérique;**
- **et sur la conception de la plateforme de service y compris l'intégration des ressources et la coopération des services.**

Resumen: Este artículo describe el Servicio Integrado de Información de la Biblioteca Digital Nacional Agrícola (DLIIS, sus siglas en inglés), una plataforma de servicios de información orientada hacia los usuarios que prestan servicios a los sectores agrícola y tecnológico de China. Dicho servicio integra integralmente un gran volumen de contenido distribuido y de recursos de información y técnicas con un mecanismo para mejorar la colaboración de las organizaciones participantes. La plataforma integrada, que forma parte del desarrollo de la biblioteca digital, sintetiza las múltiples funciones de los servicios de la biblioteca y da a los usuarios la oportunidad de tener acceso a la información cuándo y cómo lo desean. El artículo recalca el diseño de Sistema Nacional Agrícola de DLIIS y la plataforma de servicios integrados de información mediante una discusión de los siguientes aspectos:

- **el establecimiento de un Sistema Nacional Agrícola de DLIIS orientado hacia los usuarios;**
- **el diseño de una arquitectura de red de cuatro niveles (nivel interactivo con los usuarios, nivel de aplicación del sistema, nivel de integración de recursos distribuidos y nivel de apoyo técnico) para DLIIS según la teoría de niveles de bibliotecas digitales; y**
- **el diseño de la plataforma de servicios que incluye la integración de recursos y la colaboración entre servicios.**

Digital library (DL) is a new kind of information management system based on the modern idea of integrated services. This new resource utilization management system in a networked environment has been a major part of the national information structure (Hu, 2005). The DL system is based on a rational organization that optimizes massive and distributed data resources. It also combines traditional library science theory and practice. The first stage of development of the DL emphasizes digitizing of information resources while the second stage focuses on integration of the distributed and heterogeneous information resources and different services modes.

Digital Library Information Integrated Service (DLIIS) is a new user-oriented information service platform that seamlessly integrates resources and techniques with the cooperation of organizations. The DLIIS platform synthesizes the multiple functions of library services and gives the users an opportunity to access the topic they want any time at any place. DLIIS has been a key development that accommodates the new characteristics of book and information service institutions under digital networked environments and is driving efforts to improve access to agricultural knowledge.

The American Institute for Scientific Information (ISI) product-ISI Web of Knowledge is a knowledge-based intelligent research integrated system platform of academic information resources system. It provides access to the world's leading citation databases, including powerful

cited reference searching, analytics tools and over 100 years of comprehensive back file and citation data.

Integrated information systems are found in different specialized information research agencies and data processing suppliers in China. These provide information services at different levels of integration, such as the China Tsinghua Tongfang corporation products — Uniform Search Platform (USP), The National Science Library of Chinese Academy of Sciences (NSLC), the China Academic Library and Information System (CALIS), the National Science and Technology Library (NSTL), and the National Agricultural Information System (NAIS) of the Chinese Academy of Agricultural Sciences (CAAS). All of these initiatives are clear demonstrations of how research libraries, in partnership with other institutions, have began to shift into an increasingly coordinated program of information resources management as they seek to establish the next generation of web-based integrated library systems.

As is well-known, China is a large agricultural country where there are many institutions or organizations connected with agriculture. China has a large professional and technical workforce involved in agricultural research, education and extension, with extensive information support services at four levels (national, provincial, prefectural and county) and in five sectors (research, education, administration, production, and agribusiness). But there is still no comprehensive digital information library specialized in agricultural research and technology. The objective of this study is to establish a user-oriented DLIIS system in China based on the digital library model that will combine the achievements of Chinese agricultural and scientific research institutions with the theory and practices of integrated service both at home and abroad. The National Agricultural DLIIS System will be a coalition of agricultural research digital libraries. It consists of a main library (the agriculture library of the Chinese Academy of Agricultural Sciences) and more than 30 branch libraries of different provincial agricultural information research institutions in different provinces in China. The DLIIS will also co-operate with selected agricultural information research institutes and a number of specialized branch libraries that will provide specialized information resources and services. The system would promote sharing among the members of the coalition and would allow for continuous development of agricultural research digital libraries thereby fulfilling the increasing demands for innovation of China's agricultural sciences and technology.

Research Analysis of Agriculture Libraries in China

China has three main information systems — public libraries, academic libraries, and scientific and technical libraries. Agricultural information research and services institutions or libraries are an important part of scientific and technical libraries. According to the statistics, there are 59 national level agricultural research institutions, and 464 provincial level agricultural research institutions distributed all over the country. Every agricultural research institution has their research emphases and corresponding information services are necessary.

The agricultural library, supported by the Agricultural Information Institution (AII) of the Chinese Academy of Agricultural Sciences (CAAS), is the largest of its kind in Asia. Its information network center networks major agricultural research institutions both in China and abroad and plays a leading role in the agricultural information system. A national agricultural information system (NAIS) has been established to provide more effective information services for its users. It is a united platform that integrates the digital and networked information resources within the AII. The integrated digital collections are organized into four categories: CD-ROMs, e-publications, special collections and subject portal projects. Included are databases, electronic journals, books, digitized rare books, portals and subject web directories (Pan and Pi, 2006). Through this united platform, users may easily search for the needed information by subject, title, keyword, and so forth. The platform also allows users to search multiple databases concurrently.

The inception of the Agricultural Scientific and Technological Fundamental Databases Project is an example of cooperation. The aim of the project is to establish a series of scientific and technological fundamental databases to promote the development of digital agricultural information resources and information sharing. The project has organized 100 information professionals and staff from 30 provinces to collaborate on construction of databases. In all, 170 databases in agricultural disciplines have been built for the project. The disciplines covered include genetics, breeding, plant protection, crop science, soil and fertilizers, irrigation and drainage, animal husbandry and veterinary science, agricultural resources, environmental protection, production and economic management. Almost 2 million records can be browsed at the website. The partners at provincial level have their own special database such as Sichuan province sub-centers of the agricultural scientific and technology fundamental databases. The same sub-centers are located in Guangdong province and Liaoning province (The National Agricultural Data Sharing Center, 2009), as well as in other areas.

The Description of the National Agricultural DLIIS System

The National Agricultural Digital Library Information Integrated Service System has the following characteristics:

- **Type of organization** – The system consists of a central library (the agricultural library, supported by AII of CAAS) and 30 branch libraries located in provincial academies of agricultural sciences; the system also cooperates with selected agricultural research institutes and a number of specialized branch libraries which

provide specialized agricultural information resources and services.

- **Technical system** – The software requirements include interoperability, integration framework, portal and platform, middleware technology and web service construction.
- **Type of collections** – The resources include library collections, introduced databases, self service resources to build collections, subject portal, agricultural databases, agricultural expert systems, and other agricultural information research products.
- **System functions** – The system functions include sharing both resources and services; providing users an integrated environment; providing librarians methods to work interactively; overcoming the contradictions between scattered resources and the users' domain specificity; and providing a window to all Chinese agricultural researchers.

The Basic Architecture of the National Agricultural DLIIS System

The basic framework of the National Agricultural DLIIS System is derived from study and investment of the general frameworks in education and in the science and technology research field. The information system of stratification theory is applied to design the four layers of the integrated system and includes essentials such as expression style of the information resources, application area, and so on. Each layer serves an independent function, packaging the heterogeneous internal information, and providing the interface and services to connect to the external environment (Wan and Yi, 2007). System structure and content are assigned to the layer and accommodated technically.

The four layers framework of the National Agricultural DLIIS System are as follows:

Layer 1: User layer of integration services

Layer 2: Application layer of integration systems

Layer 3: Integrated layer of digital distributed resources

Layer 4: Foundation layer of information integration services and support system

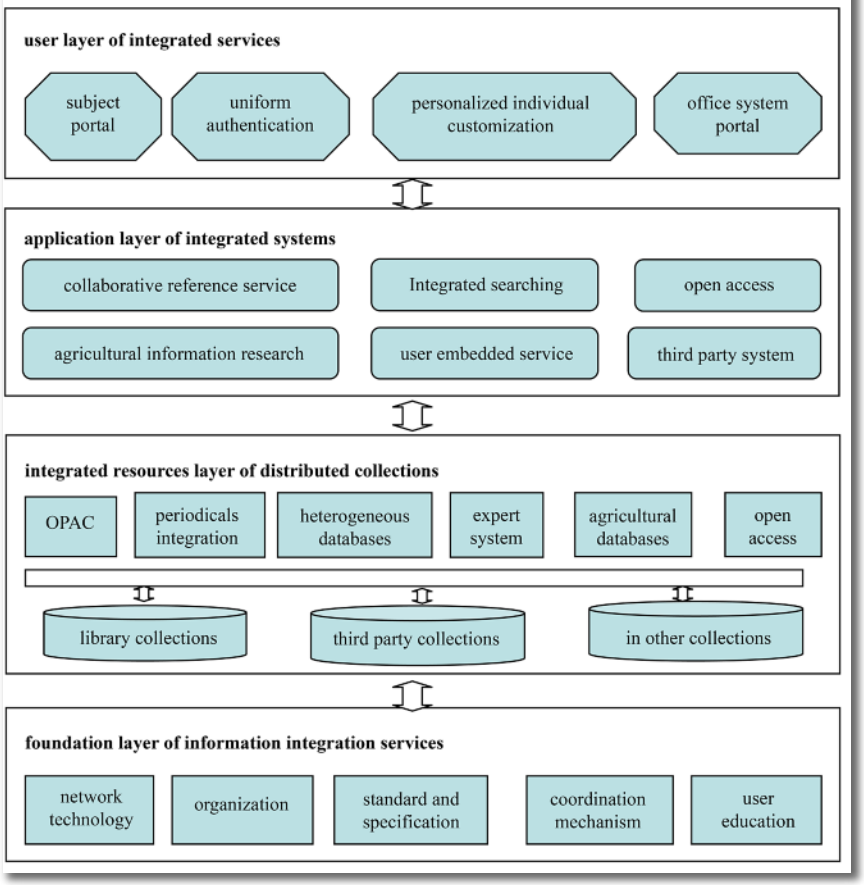
The system basic framework is shown in Figure 1.

The user layer of integrated services is the user interface, which provides uniform authentication, service applications including subject or knowledge portals, personalized or individual customization services and a mechanism for information release. All the users' needs are taken into account. One of the special functions of the user layer is that the formats of the interface can be adjusted according to the resource categories, frameworks and acquisition.

The application layer of the integrated system consists of software and application systems that support the technical contents, provide options for different integrated service methods and helps users decide what search options they want. The various applications in this layer manage the data or information processes including storage, organization, release, management, retrieval, and submission. As a middle layer between the user layer and the resources layer, the application layer plays a surrogate role.

The application layer is designed to be interoperable with distributed heterogeneous data accessible through a standard data interface. The Application layer integrates the various retrieval interfaces. It provides users a uniform window to enter the complicated system by a simple registration. It can translate research queries into the retrieval language that corresponds to the various resources through a procedure called transparent operation. Service methods included in the Application layer are open access services, cross-database searching, the

Figure 1 – The basic framework of the National Agricultural DLIIS System

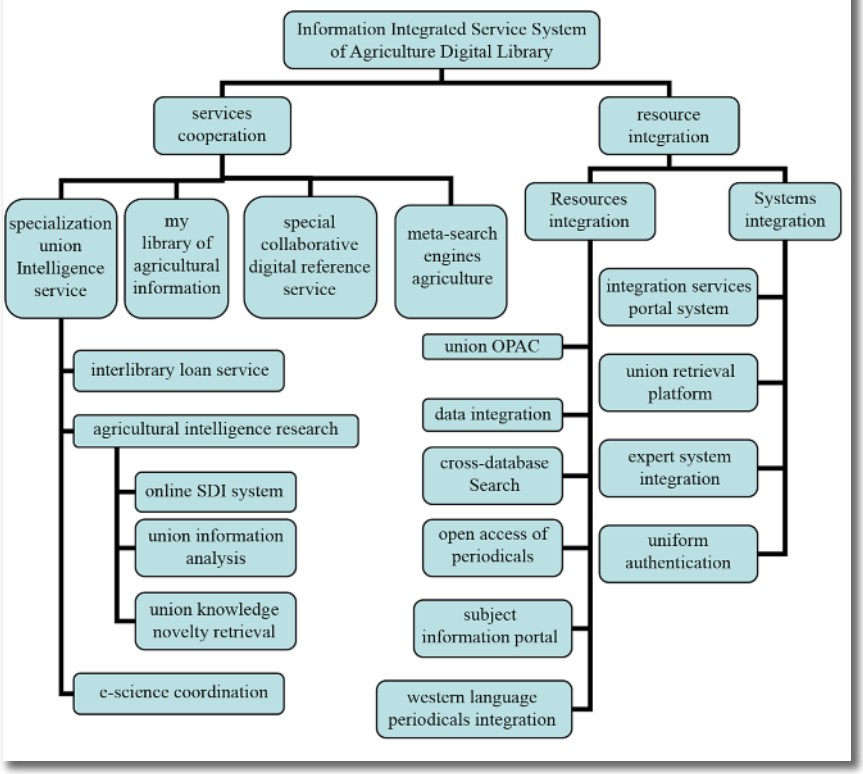


union catalog for journals, and the digitization combined reference consultation.

The Resources layer integrates heterogeneous collections distributed in more than 30 institutional repositories and connects the resources related in space and time into a multidimensional mesh structure storage system. The layer provides users with a uniform organic whole that is seamlessly linked to its heterogeneous parts. Service methods incorporated in the Resources layer are the abstracts database, full-text databases, citation databases, resource navigation databases, online public access catalogue, the special expert system, integrated periodicals and the integrated agricultural research basic data, among others.

The foundation layer of information integration services ensures the basic conditions necessary to keep the system running efficiently. The foundation layer environment contains the network hardware guarantee, the web services-based software, web servers, coordination mechanism, standards and specifications, user education, and related functions.

Figure 2 – Structure and contents of the National Agricultural DLIIS System



Structure and contents of the National Agricultural DLIIS System

The design of the platform of the National Agricultural DLIIS System is based on the principles of cooperative construction, integration and sharing, characteristic of digital libraries at home and abroad. Resources, techniques, service functions, and cooperation are compacted in two platforms — integration services and services cooperation-each having their integrated service functions and contents. There are complementarities between the two platforms in contents and information. The services cooperation platform is based on resources integration, where the resources are rationally optimized and reintegrated through the service platform. The advantage of the special Chinese system platform is that information and knowledge flow stimulated by the constant renewal of services. In this way all of the service providers and the agricultural research institutions benefit with co-construction of knowledge and information, integration services and sharing. The structure and contents of the National Agricultural DLIIS System are shown in Figure 2.

The resource integration service platform provides integrated services through resource integration. The contents include several important integrated service methods, such as union OPAC, western language journals integration, agricultural data integration, subject information portal and the union retrieval platform. The services cooperation platform provides integrated services through the synthesis and provision of different service functions, such as the personalized service system "mylibrary", collaborative agricultural digital reference services, consultation and analysis of union agricultural information and the new embedded service.

The fundamental emphasis of the resource integration service platform is the sequential clustering and availability of the resources through cross referencing or linking in an interoperable, multi-level and cross-platform manner. The key emphasis of the services cooperation platform is on librarian-initiated services and organizational cooperation based on integrated resources and on the amalgamation and extension of the various service functions.

Summary

This paper describes a project to create the National Agricultural Digital Library Information Integrated Service (DLIIS), a user-oriented information platform serving a wide range of users in the agricultural sector in China. The paper discusses the overall framework, architecture and operating platforms of the national system that seamlessly integrates diverse, heterogeneous content produced by many agricultural research organizations throughout the country and allows for its re-use

in multiple output formats. The system is designed to provide tools, applications and specialized content for both information professionals and for end users at all levels in the agricultural science and technology sectors. The study discusses how such a system, based on the principles of cooperation and interoperability, provides benefits to intermediaries and end users alike and concludes that the DLIIS is an important contribution to helping China fulfill its requirement for agricultural science and technology innovation.

References

Hu, Changping, 2005. "User-oriented resource integration and service platform strategies". *Journal of Library Science in China*, 31(2):5-9.

Pan, Shunchun, and Pi Jiezheng, 2006. "Agricultural digital information resources construction in China". *Quarterly Bulletin of International Association of Agricultural Information Specialists (IAALD)*, 51 (2):89-94.

***The National Agricultural Data Sharing Center*, [Online] available: http://www.agridata.cn/homepage/ch_intro.asp?r=1.**

Wan, Lancheng and Ao Yi, 2007. *Digital Library Technique-Information Integration and Information Retrieval*. Beijing: National Defense Industry Press, 96 p.

Contact Information

Wenfei Zhuo
Agricultural Economy and Information Research Center
Henan Academy of Agricultural Sciences
China
kfszhuo@sina.com, kjcankao@126.com

Shuchun Pan
Agricultural Information Institute of the Chinese Academy of Agricultural Sciences
Beijing
China

Zongfang Xie
Agricultural Economy and Information Research Center
Henan Academy of Agricultural Sciences
China

Enhancing Communication of Agricultural Research Outputs Using Information and Communication Technologies

Philip Edge, Stephen Rudgard, Ajit Maru and Enrica Porcari

Abstract: Information and communication technologies (ICTs) have transformed the ways in which information and knowledge arising from agricultural research can be shared. Yet the formal outputs of research are often not openly available, and they are not digitally disseminated and communicated effectively and coherently. In addition, some developing countries have been slow to benefit from these new digital opportunities. A group of international and regional organizations established the CIARD (Coherence in Information for Agricultural Research for Development) initiative in 2008 to address issues of access, coherence and capacity in agricultural research systems. CIARD works through the three strategic strands of Advocacy, Capacity Building, and Content Management. The CIARD Manifesto, Checklist, and Pathways, together comprise a strategic and practical framework which can be used and customized by organizations at any level. The website presents the elements of the framework, describes the partnership, and also provides access to the CIARD-RING, a public-domain global registry of web-based sources and services that give access to information and knowledge. The three main strands of action will continue to be pursued to enhance CIARD's influence and increase access to knowledge worldwide through face-to-face meetings and online collaboration to foster sharing of experiences.

Resumé: Les technologies de l'information et de la communication (TIC) ont transformé la façon dont l'information et la connaissance issues de la recherche agricole peuvent être partagées. Mais les résultats formels de la recherche ne sont pas toujours du domaine public, et ne sont pas diffusés en format numérique ni communiqués de manière efficace et cohérente. De plus, certains pays en développement ont tardé à bénéficier de ces nouvelles possibilités numériques. Un groupe d'organisations régionales et internationales a lancé en 2008 l'initiative CIARD (Cohérence de l'information sur la recherche agricole pour le développement) dans l'objectif d'apporter une solution aux problèmes d'accès, de cohérence et de capacités dans les systèmes de recherche agricole. Le CIARD a trois axes stratégiques d'intervention: Plaidoyer, Renforcement des capacités et Gestion de contenu. Le manifeste, les recommandations et les itinéraires du CIARD constituent ensemble un cadre pratique et stratégique qui peut être utilisé et personnalisé par les organisations à tous les niveaux. Le site web présente les éléments du cadre, décrit les partenariats et donne également accès au CIARD-RING, un registre mondial du domaine public de services et de sources sur le Web qui permet d'accéder à l'information et à la connaissance. Les trois grands axes d'intervention continueront d'être privilégiés pour renforcer l'influence du CIARD et l'accès à la connaissance dans le monde, grâce à des rencontres et à la collaboration en ligne qui faciliteront le partage des expériences.

Resumen: Las tecnologías de la información y comunicaciones (TIC) han transformado la forma en que se comparte la información y conocimientos que emanan de la investigación agraria. Sin embargo, los productos formales de dicha investigación no siempre se encuentran disponibles ampliamente; tampoco se difunden digitalmente o comunican de manera eficaz y coherente. Por otra parte, el proceso de adopción de estas nuevas oportunidades digitales ha sido bastante lento en algunos países en desarrollo. En 2008 se reunió un grupo de organizaciones internacionales y regionales para crear la iniciativa CIARD (Coherencia en la información para la investigación agraria para el desarrollo) a objeto de abordar los problemas de acceso, coherencia y capacidad de los sistemas de investigación agraria. CIARD se rige por tres líneas estratégicas: Promoción, Desarrollo de capacidades, y Gestión de contenidos. El Manifiesto, Lista de verificación y Rutas para CIARD en su conjunto constituyen el marco estratégico y práctico que las organizaciones, a cualquier nivel, pueden utilizar y adaptar. El sitio web presenta los elementos de dicho marco de trabajo, describe la asociación y ofrece acceso a CIARD RING, un registro global de dominio público de fuentes y servicios basados en la web que permiten acceder a información y conocimientos. Las tres principales líneas de acción continuarán desarrollándose a fin de aumentar la influencia de CIARD y el acceso global al conocimiento a través de coloquios presenciales y la colaboración en línea para fomentar el intercambio de experiencias.

Rationale

Researchers and research systems everywhere are witnessing major changes in the ways in which the results of research, and the ensuing applications and impacts, are being made accessible and communicated. The processes by which knowledge, information and data are generated and shared are being transformed and reinvented by the application of information and communication technologies (ICTs). These changes are providing opportunities for agricultural researchers worldwide to interact and participate together in the development and usage of their research. Further, the whole information chain, from researcher to extension worker to farmer and back again, is being transformed by the use of ICTs.

The last decade has seen rapid growth in the open availability on the Internet of research information. The Open Access (OA) movement in developed countries has primarily focused upon making the formal outputs of research (research papers) freely and openly available on the Internet, either through open access publishers or by deposit into institutional or subject-based repositories. The technologies required to build and operate

open access repositories have developed at low cost. Aided by this factor the great potential for developing country institutions and networks to participate in the digital knowledge economy has become apparent. Open repositories in particular bring into focus the opportunities to capture and share a range of types of research content, including theses, data, images, researcher profiles, and so on. In addition, the more recent rapid spread of Web 2.0 tools and related communication activities, has opened up low cost routes to the capture and dissemination of research outputs in more informal ways. In this new environment it has become increasingly clear that research outputs that are not captured, disseminated and communicated in an effective way are largely invisible. Invisibility means that the investment in that research has been lost.

Because of low investment, in both technical infrastructure and agricultural research, some developing countries have been slow to benefit from this new, digital, open environment. Further, the development of digital information systems needs to conform to international standards and methodologies. A lack of technical "coherence" undermines efforts to combine and share information and knowledge between countries and continents. The other dimension of coherence lies in the arena of effective actions taken by organizations that can align their efforts to develop information collections and services that are interconnected and more accessible, avoiding competition between systems and initiatives. Although each has its own priorities and constraints, all subscribe to common approaches.

It is within this context of the rapidly changing landscape of research communication that a group of major international organizations, all with a long-standing role in enhancing the sharing of information and knowledge arising from agricultural research, came together to combine their experience and to address the issues of coherence and capacity in developing country research systems. FAO, GFAR, CIRAD, DFID and the CGIAR, organized two expert consultations on "International Information Systems for Agricultural Science and Technology" (IISAST) in 2005 and 2007. These events brought together representatives of the key regional institutions and other major actors, who produced a series of key recommendations to enhance various aspects of coherence.

However, the partners were unable to develop the basis of a profile for the initiative that all parties could buy into. The principal partners met again in a small landmark meeting in January 2008 at the French Ministry of Foreign Affairs to devise a shared identity, and they settled on the name CIARD (Coherence in Information for Agricultural Research for Development) for the initiative, and the following vision: "To make public domain agricultural research information and knowledge truly accessible to all", with the aim of enabling organisations and individuals that create or possess public agricultural knowledge to disseminate it more effectively. The principal elements of the CIARD initiative were finalised with the wider group of partners in 2008.

Public knowledge and research has a limited impact on agricultural and rural development and natural resources management, because most of these outputs are not easily or widely accessible. CIARD aims to change this situation with the message that the new ICTs provide many opportunities for information to be handled and presented differently and economically. The CIARD Partners coordinate their efforts, promote common formats for information sharing and exchange, and adopt open information systems approaches. However, CIARD is not just about technology. It actually addresses the ways that technologies are used. This includes building and improving information systems, empowering the institutions and people using them with a framework and a set of tools that open access to their resources. So, while every institution or individual exists in their own environment, CIARD encourages them to interconnect and work together in ways that complement each other. In this way, CIARD addresses the fundamental issues involved in making local, national, regional and global information systems available and accessible.

In order to collaboratively develop common standards, share knowledge, and contribute to effective and coherent institutional and individual approaches in agricultural science and technology information, CIARD has focused activity in three priority areas:

- Make content accessible: with open content, open systems and common international standards;
- Develop capacities: by empowering individuals with awareness and skills and encouraging institutions to be self-sufficient through ownership of their information;
- Advocate better investments: with policies that enable easier access to information, coordinated approaches, and using evidence of benefits.

The current situation

The CIARD initiative is being developed in several different ways.

CIARD is working to achieve its' aims through three strategic strands in the form of Task Forces. The Content Management Task Force addresses the development and promotion of: methodologies and tools for Open Access repositories; common standards and protocols for agricultural information exchange; and common services to access partners' distributed information. The Capacity Building Task Force addresses issues related to capacity development of institutions and individuals that wish to align with the CIARD Manifesto, including training for individuals and development of training resources. The Advocacy Task Force is developing a sustained effort to increase awareness of and support for the CIARD Manifesto amongst key stakeholders at policy and institutional

level. These Task Forces use virtual collaboration facilities, and they meet as necessary at international events to progress policy and implementation initiatives, and to promote the CIARD concepts.

A series of regional expert consultations driven by the Advocacy Task Force were organized to gather inputs from more than 150 experts from almost 70 countries and a range of regional and international organizations, who are directly involved with the communication of research outputs. These consultations allowed real experiences to influence and guide the nature of the concepts being developed in the CIARD Manifesto, the Checklist of Good Practices (which illustrates to institutional and individual users the key areas of activity which, if followed, will enable them to participate in the international digital information economy), and the Pathways (see below). The consultations led to a strong endorsement of the CIARD agenda, with the contributions resulting in a major revision of the CIARD Manifesto, and the Checklist of Good Practices was reformulated into its' current shortened form (see below). As a result of this, the CIARD website was entirely reconsidered and the layout enhanced with appropriate alterations to the main elements of text. This process of consultation has also provided momentum to a plan to promote further the CIARD programme by:

- introduction of the CIARD Manifesto and Values to research communication specialists in Europe and North America through major international/regional meetings, areas not covered in the 2009 consultations;
- development and dissemination of the CIARD advocacy tools in as many languages as possible;
- increase in efforts to gather evidence of development outcomes from national and local levels from institutions following the CIARD approach;
- continuing enrolment of institutions into the CIARD community.

The explanation and the routes for implementation, of the CIARD agenda as a whole are delivered through the CIARD website (http://www.ciard.net/) and associated print materials. These show the CIARD Manifesto and the Checklist of Good Practices. Then, a collection of Pathways is available (currently 16), each one explaining how the user can address a specific area of activity (such as creating a policy framework to enable an institution to develop a repository) necessary for the support of the CIARD agenda. The Pathways cover the development of institutional readiness, the collection and preservation of information and data (including building a repository), and making content accessible on the web. All Pathways address the CIARD agenda of making information available, accessible and applicable in the digital environment.

The website also provides the CIARD RING (Routemap to Information Nodes and Gateways), a global registry of web-based services that will give access to any kind of information sources pertaining to agricultural research for development (ARD). The RING is the principal tool created through the CIARD initiative to allow information providers to register their services in various categories and so facilitate the discovery of sources of agriculture-related information across the world. The RING aims to provide an infrastructure to improve the accessibility of the outputs of agricultural research and of information relevant to CIARD management.

Regional Expert Consultations

Africa (Accra)
Asia and the Pacific (Hyderabad)
Latin America (Santiago)
Middle East (Cairo)
Russian Confederation (Moscow)

Lessons learned from experience, and promising trends

The 'Manifesto' and 'Checklist of Good Practices'

The CIARD Manifesto (http://www.ciard.net/ciard-manifesto) presents a context which explains the issues that the initiative aims to address, and a set of values which the initiative embraces. The Checklist of Good Practices (http://www.ciard.net/checklist) takes these issues a step further.

A conceptual model (Figure 1), which emerged originally from the first IISAST Expert Consultation, categorizes the complex "reality" experienced by institutions

Figure 1 – Conceptual model from the first IISAST Expert Consultation

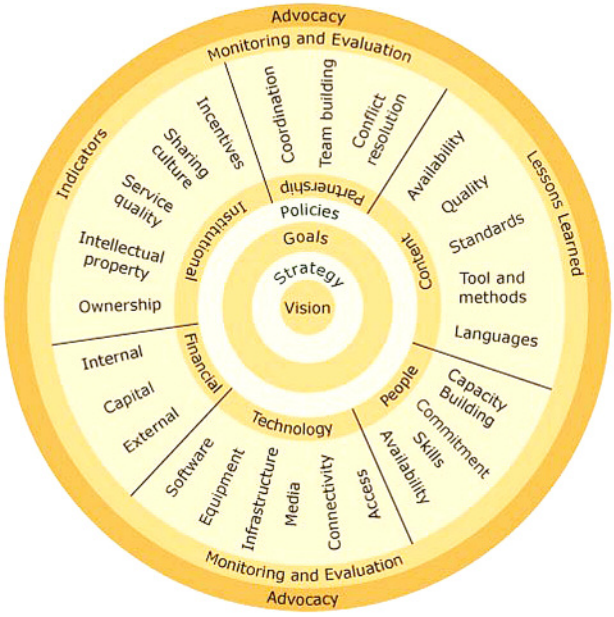


Figure 2 – CIARD Checklist of Good Practices

Developing Institutional Readiness

1. Introduce and gain support for the CIARD Manifesto and Values in your institution.
2. Have your institution recognised as a CIARD partner.
3. Adopt a formal institutional information / communication strategy.
4. Develop the capacities of your institution to achieve the CIARD Checklist.
5. Develop national/local partner networks to share resources and skills.

Increasing the Availability, Accessibility and Applicability of Research Outputs

6. Ensure that your research outputs are available digitally
7. Develop institutional or thematic repositories of your outputs as open archives.
8. Use international metadata standards, data exchange protocols and agricultural vocabularies and thesauri.
9. Develop a clearly defined licensing policy for your outputs.
10. Optimise the structure and the content of your web sites for search engines.
11. Share your metadata by participating in international information systems.
12. Use 'social networking' media and applications to share your outputs.
13. Build formal and informal networks to repackage your outputs.

and national networks which had already embarked upon major digital information projects. The model maps the major elements that are required for successful development of agricultural information systems in institutions and networks.

The CIARD Checklist represents a set of items covering the elements of the above model through which organisations, research systems and individuals can work towards achievement of the CIARD Manifesto. It is not a set of requirements. There are many different ways in which the Checklist agenda can be achieved, according to the specific scope of an organisation's work. So everyone will be able to develop their own route through the Checklist. An organisation or system can assess itself against the Checklist to reveal its status, and then can prioritise and stage actions over time to achieve as many of the Checklist items as are appropriate to local needs. Organizations are encouraged to follow this process on the CIARD website. Implementation of many of the Checklist items will depend on capacity building in important techniques for key actors. The Checklist actions address a range of stakeholders, setting out approaches that will ensure that research outcomes are more likely to be sustainable.

The current status of the Checklist of Good Practices is shown in Figure 2. As is shown, it presents two areas of activity — Developing Institutional Readiness (critical to the effective development of necessary change within an organization), and Increasing the Availability, Accessibility and Applicability of Research Outputs (referred to as the 3 A's).

The current status of the Checklist has been achieved over a period of almost two years. This process began by using the increasing body of knowledge available on best practice in digital information management, which was refined by, and put into context by, the breadth and depth of experience of the CIARD Partners. Most importantly, the list has been further refined by key inputs from the regional expert consultations referred to above. This evolution will continue, wherever necessary, in response to input from users providing insights into the real issues they face in their local organizations and networks.

The CIARD Pathways

In addition to the Checklist of Good Practices, a range of targeted CIARD Pathways have been developed showing the ways in which the Checklist actions can be achieved.

List of Acronyms

CGIAR	Consultative Group for International Agricultural Research
CIARD	Coherence in Information for Agricultural and Information Development
CIARD-RING	CIARD Routemap to Information Nodes and Gateways
CIRAD	Agricultural Research for Development (France)
DFID	Department for International Development (United Kingdom)
FAO	Food and Agricultural Organization of the United Nations
GFAR	Global Forum for Agricultural Research
ICTs	Information and Communication Technologies
IISAST	International Information Systems for Agricultural Science and Technology
IMARK	Information Management Resource Kit
OA	Open Access

Figure 3 – CIARD Pathways

Group 1: Developing Institutional Readiness

- Advocate the benefits of the digital accessibility of content
- License Content to Encourage Use and Re-use
- Work with Publishers who have Flexible Policies on Open Access

Group 2: Collection and Preservation

- Digitize older outputs, to be 'born again' digital
- Put in place institutional policies that enable the sustainable development of a repository
- Preservation of digital documents and data
- Develop a Repository for Digital Content, and Repository Development Case Studies:
  - A new challenge: an Agricultural Knowledge Repository for Thailand
  - E-LIS – E-prints in Library and Information Science
  - NARIMS – National Agriculture Research Information Management System
  - WaY (Wageningen Yield – The Repository of Wageningen University and Research Centre
- Maintaining web links and managing broken links

Group 3: Making Content Widely Accessible on the Web

- Disseminating research outputs — international databases
- Making a website's content visible on the Web
- Set up added value services that query across platforms
- Publish and promote outputs with newsfeeds
- Using social media to communicate research outputs
- Using video to communicate research outputs
- Using Web 2.0 solutions for your Website
- Analyse how your websites are being used, and put this knowledge to use

Figure 4 – Introduction to Pathway Group 1: *Developing Institutional Readiness*

Intended audience:
Research managers, scientists, information professionals

The CIARD Pathways provide an introduction to the many ways in which research outputs can be made more available, accessible and applicable to stakeholders who will derive benefits from this knowledge.

For the administrator/policymaker, research director, individual researcher and librarian alike the Internet, the open access movement, and the development of digital repositories have created new possibilities for enhancing the 'visibility' of research outputs and have greatly increased the potential audience for them.

Within the CIARD Pathways there are several which focus particularly on areas of change in both policy frameworks and individual researcher behaviour which will help to prepare an institution to stimulate the dissemination of research outputs. These Pathways are:

1. Advocate the benefits of the digital accessibility of content
2. License content to encourage use and re-use
3. Work with publishers who have flexible policies on Open Access

This group of Pathways introduce:

- Ways in which digital information management can be championed and advocated within an institution or network in order to gain the support of all relevant stakeholders.
- The types of policy that will govern the effective development and continuing management of a digital repository — planning and resourcing, what should be deposited and how, the assessment of usage and success, and so on.
- How researchers and research groups can ensure the maximum visibility for their outputs by publication in open access journals, or working with publishers who allow maximum flexibility for the author by allowing them to post their work freely on web sites or in repositories.
- How researchers and institutions can create policies that encourage the use and reuse of their research outputs within a licensing framework that, while stimulating the dissemination and use of the outputs, also protects their integrity and the rights of the author and his or her institution.

These Pathways show how policy frameworks, combined with advocacy and the commitment of individuals, will help research outputs to become more 'visible' and widely disseminated and hence raise the reputation of both institutions and individual researchers.

The Pathways (currently 16 in number) provide detailed practical guidance in specific areas of activity. It is unlikely that any single network or organization or individual would need to follow all Pathways. The list of current Pathways are shown in Figure 3. Both collectively and individually the Pathways have been developed through a process of bringing together best practice information from the current corpus of knowledge on digital information management. This body of best practice is steadily being established by the activities of researchers and information management specialists worldwide. They have then been further refined through a process of input from workshops and consultation with experts with broad experience in the fields covered.

All CIARD Pathways, individually and as a list, are completely open to further development responding to input from the real experiences of users in all regions. Further, the Pathways will be produced in several languages to extend their dissemination and usage.

In their current state of development the Pathways have been organized into three Groups, as shown in Figure 3, which relate to the organization of the Checklist of Good Practices. Group 1 concerns 'Institutional Readiness'. Groups 2 and 3 cover 'Collection and Preservation' and 'Making Content Widely Accessible on the Web'. As a whole they aim to deliver the 3 A's of information and knowledge sharing (Availability, Accessibility and Applicability).

To help to put the Pathways into context for potential users each Group is described by an Introduction which explains the intended audience for those Pathways and what the user will learn from them. The 'Introduction to Group 1' is shown as an example in Figure 4.

Each individual Pathway is presented in a standard form, starting with a brief introduction to the context in which the Pathway has been developed, what the user needs to know in terms of background, and what they will need to do to if they wish to follow the Pathway into further detail. References and links to more detailed information resources are provided. The purpose of each Pathway is to introduce the user to that particular area of knowledge, explain its' relevance, and then to enable the user to pursue the subject in further depth if that is required. An abbreviated example of a Pathway is shown in Figure 5.

Principles and guidelines for assessing, designing and implementing interventions

The development so far of CIARD interventions (ranging from workshops assessing content management tools, to establishing documentation to be used in advocacy campaigns) has been critically dependent on three elements:

- The use of best practice, openly available, guidance on approaches to creating change in information and knowledge management for individuals and organizations.
- The breadth and depth of experience of the CIARD Partners.
- An ongoing, iterative process of gaining feedback at all levels of the programme (from Manifesto and Values to the specifics of a particular Pathway).

The outcome of this process can be seen in the Checklist of Good Practices and the development of the Pathways both described above. In both cases international best practice across a number of different areas of technical and organizational study has been brought together, adapted and redefined so as to be meaningful to the constituencies which are to be the chief beneficiaries of the CIARD programme.

On the CIARD website all concepts and documentation are open to detailed comment and feedback by anyone registered on the site. In this way the CIARD initiative is open to constant development based on the real experiences of users, whether contributed through meetings and events or through comments on the website.

Cross-cutting issues and key enablers

The new environment of research communication in the digital age is having impacts now on everyone involved, directly or indirectly, with agricultural research, whether or not they are fully aware of it. The potential benefits for developing countries by involvement in these systems are very great. The open exchange of research information, and the collaborative involvement with research problems across national and regional boundaries, has never been so enabled. The technical infrastructure required is available to all at relatively low cost.

What is often holding back the active involvement of individuals and institutions is the lack of understanding of the potential to be tapped. This lack of understanding can occur at all levels, from the researcher, to the institutional manager, to the policy maker. It is this lack of understanding in particular that CIARD aims to address. At the spearhead of CIARD activity is the strategic strand of Advocacy. Through presentations at conferences, workshops and expert consultations, and marketing campaigns, CIARD is working to get across the messages of 'digitization', 'openness' and 'coherence'.

The regional expert consultations referred to in Section 2 above, apart from producing detailed input into the formulation of the CIARD Manifesto and Vision, Checklist of Good Practices, Pathways, and their presentation in the CIARD website, also identified a number of common concerns and challenges, such as:

- How to identify the resources, both technical and in other skills, needed to make the changes necessary to achieve the Checklist.
- How to persuade and mobilise the three key groups of stakeholders in institutions, namely researchers, policy-makers and information managers, into communicating their research outputs.

- The need for CIARD to harmonize with the agendas of national and sub-regional organizations (and vice versa) attempting to address the same issues.
- The richness of the CIARD agenda provides a context within which institutional and national developments can take place. But how can the best alignments be created?
- The challenges of turning the Checklist and Pathways into action.

There are no simple solutions to these challenges. They are experienced in much the same way in developed countries, but the difficulties are more acute, perhaps harder to solve, in the developing world. The CIARD programme, through its' Advocacy, Capacity-Building and Content Management strategic strands is able to give insights into the solutions.

It is becoming clear that a number of features of the CIARD programme, as it has evolved so far, are central to its' success and will help to maintain its' further development.

1. CIARD is a partnership and not an organization or a legal entity. It is a partnership of like-minded organi-

Figure 5 – (Abbreviated) Pathway: *Advocate the benefits of the digital accessibility of content*

Intended audience:
Research managers, scientists, information professionals

Technical and cultural barriers slow the acceptance of change by both individuals and institutions. You will need to address these barriers in your institution and with individual colleagues if change is to happen. This Pathway shows you how to start to advocate for the digital accessibility of research outputs. It tells you about the arguments and tools you can to use to encourage all stakeholders to work in the digital world.

What do you need to do?

1. You need to advocate for the development of supported and managed digital resources in your institution. All stakeholders (researchers, librarians, IT staff, administrators and students) need to know of the advantages that this will bring. This learning process can be stimulated through a campaign of presentations and other communication targeting the key stakeholders. The advantages to the institution and to the individual need to be made clear. But also important is the context of the bigger picture of research information nationally and internationally. A selection of the major benefits to strengthen the case for 'going digital' are shown below; see References for further help:

   For the researcher:

   - Increased visibility of research outputs and consequently visibility of the department and the institution.
     Etc.

   For the institution:

   - Increases the visibility and prestige of an institution.
     Etc.

   For the global research community:

   - Assists research collaboration through facilitating free exchange of scholarly information.
     Etc.

2. Your case will be strengthened by carrying out an information needs assessment, or by benchmarking accessibility, to show how information is currently produced and managed by your institution, and individuals within it, and the consequent benefits in moving to digital management.

3. Be clear about the strategy for digital development and for a repository. Include the following strategic elements right from the start:

   - what is the repository going to do and why (manage institutional information better, promote the quality of your institution's work, contribute to national research outputs, and so on)? Etc.

4. How do you change the culture of an institution? It is important to develop strategic plans collaboratively with all key stakeholders. An institution is driven by a combination of intellectual, emotional and political motivations. It is not easy to change hearts and minds — it takes time. This will be done through communication, persuasion, and the well managed strategic planning of developments. But the clarity of the benefits to all stakeholders will eventually win.

Examples

- Greece: 'The case of "Eureka!" the Institutional Repository of Alexander Technological Educational Institution.' (http://www.rsp.ac.uk/repos/casestudies/thessaloniki.php)
  Etc.

References

Detailed advice and information on all aspects of advocacy and making your case within an institution and to colleagues can be found at:

- EPrints (www.eprints.org)
- Repositories Support Project (RSP) — see particularly sections on 'Before you start' and 'Advocacy options'. (www.rsp.ac.uk)
- Open Access Scholarly Information Sourcebook (OASIS) (www.openoasis.org/)

zations working toward a common end. The "Coherence" refers not only to the technical reality but collaboration in an effective partnership. The founding Partners have committed their own resources to enable their continuing support for and participation in the initiative. Further, any organization can now become a CIARD Partner by registering its' interest and capabilities on the website, and can contribute to moving the initiative forward.

2. CIARD's openness to new thoughts and new thinking and the interactivity with users through its' website. CIARD is committed to maintaining this flexibility to enable it to evolve further.
3. The commitment to capacity-building, which is so central to long term success. FAO, as one of the CIARD founding Partners, is also the leader in a partnership developing the Information Management Resource Kit (IMARK — http://www.imarkgroup.org/). IMARK is an e-learning initiative, available on CD and the Internet, to train individuals and support institutions and networks in the effective management of agricultural information. IMARK consists of a suite of distance learning resources, tools and communities on information management. The many areas of synergy between IMARK and the CIARD programme means that IMARK, which is also freely available, provides a powerful platform of support for the capacity-building efforts required by CIARD.

Next steps

The initial phase of the inception of CIARD in 2008 and 2009 focused on the development of the main elements of the framework, the Manifesto, Checklist and Pathways, and the design and launch of the RING. The consultations that were organized to validate these elements doubled as promotional opportunities to raise awareness to the initiative. The second phase in 2010 looked to bring the main framework to major international events to seek confirmation that there was support amongst the international agricultural research community for taking the initiative forward, and this was achieved at the global level, in Africa, in Europe, and in the USA.

Future efforts by the principal sponsors/partners will concentrate on maturing the CIARD initiative into a broad scale international movement with widespread uptake and recognition. Promotion of the initiative will be intensified, for which an advocacy toolkit will be compiled and published, ideally in several languages, to assist those who wish to promote the initiative. Some new features will be launched on the main web platform. A Virtual Marketplace will be opened on the CIARD website that shows how the Pathways can be achieved with the help of "enablers" who assist with achievement of the Pathways, in terms of software, information management tools and methodologies, products and services that assist information discovery. New Pathways will be published aimed particularly at research scientists. A range of case studies will be documented and published to show how organizations have achieved the Checklist and used the Pathways. Finally a series of virtual and face-to-face international consultations will be organized to facilitate the development and sharing of a range of coherence standards for different types of digital information objects. In the meantime, some concrete measures of success for the CIARD initiative will be monitored in terms of the number of organizations that have registered on the main website, and the number and size of the collections of digital objects that have been registered on the RING. There will also be indirect and less tangible outcomes that arise from institutional efforts to share the outputs of research using the elements of the CIARD framework, but which are not formally registered in the CIARD domain.

References

EUR 23459 — Open Access — Opportunities and challenges — A handbook (European Commission/German Commission for UNESCO). 2008. Luxembourg: Office for Official Publications of the European Communities 144 p. DOI: 10.2777/93994 http://ec.europa.eu/research/science-society/document_library/pdf_06/open-access-handbook_en.pdf [accessed February 2011]

Kirsop B., Arunachalam S., Chan L. (2007) Access to Scientific Knowledge for Sustainable Development: Options for Developing Countries. *Ariadne*, 52 (July 2007). http://www.ariadne.ac.uk/issue52/kirsop-et-al/ [accessed February 2011]

Web 2.0. *Wikipedia* http://en.wikipedia.org/wiki/Web_2.0 [accessed February 2011]

Technical Centre for Agricultural and Rural Cooperation (CTA) et. Al. Report of the 2nd Expert Consultation International Information Systems for Agricultural Science and Technology 23-24 September 2007. http://www.iaald.org/docs/iisast2_report.pdf [accessed February 2011]

Contact Information

Philip Edge, Consultant

Stephen Rudgard
Food and Agriculture of the United Nations (FAO)
Rome
Italy
Stephen.Rudgard@fao

Ajit Maru
Global Forum for Agricultural Research (GFAR)
Rome
Italy
Ajit.Maru@fao.org

Enrica Porcari
Consultative Group for International Agricultural Research (CGIAR)
Rome
Italy
e.porcari@cgiar.org

Disseminating the Outputs of Agricultural Research in Africa

Justin Chisenga, Richard Kedemi, Joel Sam, Stephen Rudgard and Franz Martin

Abstract: **Effective access to public domain scientific knowledge is critical to a broad range of stakeholders and communities across Africa to develop and apply solutions for rural development and to improved rural livelihoods. This paper describes the development of pilot networks implemented in Ghana and Kenya to facilitate open access to public domain research information through digital repositories that make agricultural research outputs accessible to all. The networks were developed over three years drawing on the resources, tools, and technologies available from FAO through the AGRIS network and from other sources. Strong progress was made in these pilot national networks of institutional repositories, with enabling policies formulated through participatory approaches and significant quantities of full text content now accessible in digital format. The experiences and lessons learned gained by these initiatives have been fed into the development of the framework for the CIARD initiative in 2009, and these pilot networks are set to expand to national level.**

Resumé: **Un accès efficace à la connaissance scientifique du domaine public est indispensable à toute une gamme de parties prenantes et de communautés en Afrique pour pouvoir élaborer et appliquer des solutions pour le développement rural, et améliorer les moyens d'existence en milieu rural. Le présent document propose une description de réseaux pilotes mis en place au Ghana et au Kenya afin de faciliter l'accès ouvert aux informations relatives à la recherche du domaine public, grâce à des référentiels numériques qui rendent les résultats de la recherche agricole accessibles à tous. Les réseaux ont été créés en trois ans, à l'aide de ressources, d'outils et de technologies de la FAO disponibles par le biais du réseau AGRIS et d'autres sources. De progrès marquants ont été réalisés au niveau de ces réseaux pilotes nationaux de référentiels institutionnels: des politiques de facilitation ont été formulées selon une approche participative et des quantités considérables de contenu en texte intégral sont maintenant accessibles en format numérique. Les expériences et les enseignements tirés de ces activités ont largement contribué à l'élaboration du cadre de l'initiative CIARD en 2009, et ces réseaux pilotes sont appelés à s'étendre au niveau national.**

Resumen: **Existe un amplio espectro de actores y comunidades en África para los cuales se hace imprescindible contar con acceso efectivo al conocimiento científico de dominio público para desarrollar e implementar soluciones para el desarrollo rural y mejorar los medios de vida rurales. Este artículo describe el desarrollo de redes piloto implementadas en Ghana y Kenya para facilitar el acceso abierto a información de investigaciones de dominio público a través de repositorios digitales que permiten a cualquier usuario acceder a los productos de la investigación agraria. Las redes se fueron desarrollando a lo largo de un período de tres años teniendo como base recursos, herramientas y tecnologías de la FAO, disponibles a través de la red AGRIS, así como de otras fuentes. El avance más notable en estas redes nacionales piloto de repositorios institucionales se logró implementando políticas que han sido formuladas empleando enfoques participativos y una cantidad significativa de contenido de texto completo disponible actualmente en formato digital. Las experiencias que surgieron y lecciones aprendidas de estas iniciativas impulsaron la consolidación de la iniciativa CIARD en 2009 y se prevé que estas redes pilotos se extiendan y alcancen nivel nacional.**

The Problem: Limited access to research information

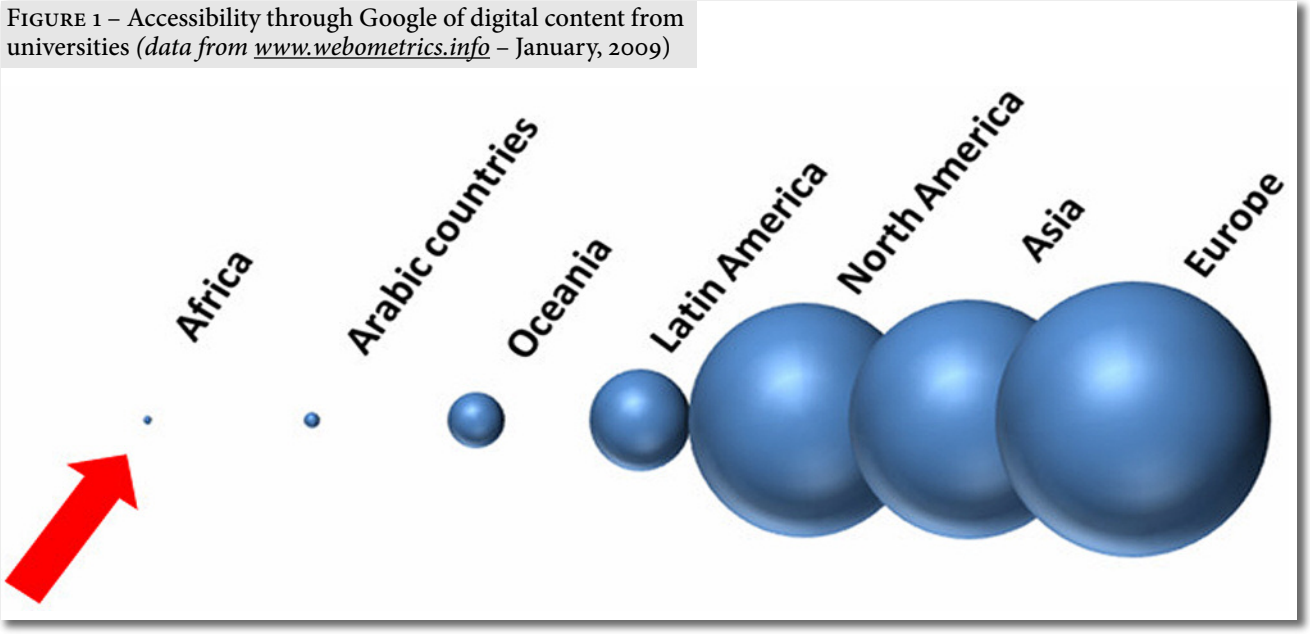
Effective access to public domain scientific knowledge is critical to a broad range of stakeholders and communities across Africa to develop and apply solutions for rural development and to improved rural livelihoods. Figure 1 reveals that there is very limited access to digital information in Africa in terms of the amount of web-based content available and accessible as compared with other regions.

This paper reports the findings of initiatives in two African countries to facilitate open access to public domain scientific and technical information on agriculture through interlinked full-text repositories at the institutional level, developed as part of a national agricultural science and technology system. The aim was to foster improved archiving, dissemination and sharing between researchers and other stakeholder groups.

These pilots were developed over three years, through a programme that was co-funded by Food and Agricultural Organization of the United Nations (FAO) and United Kingdom's Department for International Development (DFID) developed with the aim for the experience gained by the institutions to serve as the basis firstly to test approaches and secondly expand such pilots to national level. The work on the national network in Kenya started in 2005/06 and was completed in early 2009. The work in Ghana started in 2007/08, and was completed in 2010. The pilot implementations drew on the resources, tools, and technologies available from FAO and through the international AGRIS network and other sources as appropriate.

Each of the pilot initiatives established and strengthened a national coordinating centre and up to six selected partner institutions to develop open access, public domain scientific and technical information services on agriculture through interlinked full-text repositories at institutional level. This involved the development of strategies and policies for all participating institutions, while also building institutional capacity and strengthening human resources to manage the systems and resources.

Figure 1 – Accessibility through Google of digital content from universities *(data from www.webometrics.info – January, 2009)*



Access to the outputs of public research in agriculture: Kenya

The Actors

This initiative focused on the development of an electronic repository as part of a Kenyan national agricultural science and technology information system, in relation to the Strategy for Revitalizing Agriculture. The lead institution was the Kenya Agricultural Research Institute (KARI).

Five principal institutional actors in Kenya have created a shared open repository of the outputs of agricultural research through a collaborative partnership named the Kenya Agricultural Information Network (KAINet). The creation of the shared resource has been enabled by the participatory development of supportive institutional policies in each national centre, and by extensive training and roll-out of technologies that ensure coherence nationally and with international information-sharing systems. KAINet was to promote the philosophy of the open access to information approach espoused by the Coherence in Information for Agricultural Research and Development (CIARD) initiative.

Initiated in April 2006, KAINet was developed in response to demand from the national and international community to promote information exchange and access among stakeholders in the agricultural sector to support decision-making, to promote innovation in agriculture, and to improve livelihoods. KAINet addresses the national policy to build a Kenyan national agricultural science and technology information system, enshrined in the national Strategy for Revitalising Agriculture (SRA). The SRA prioritised knowledge-sharing links between the national research system and extension and other rural service providers in Kenya, as well as international information systems. At the international level, KAINet is a response from Kenya to the call to join the CIARD global movement aimed at making public domain agricultural research information and knowledge truly accessible to all.

The main stakeholders in the first phase of KAINet were five leading national institutions: the Kenya Agricultural Research Institute (KARI), the Kenya National Agricultural Research Laboratories (KARI-NARL), the Kenya Forestry Research Institute (KEFRI), the Ministry of Agriculture (MoA) and Jomo Kenya University of Agriculture and Forestry (JKUAT), although other key players in agricultural research and extension sector in Kenya were consulted at various stages. At the international level, FAO, CABI Africa and the Association for Strengthening Agricultural Research in East and Central Africa (ASARECA) provided inputs at the policy and technical level, as well as providing training.

The Activities

KAINet development occurred in five phases:

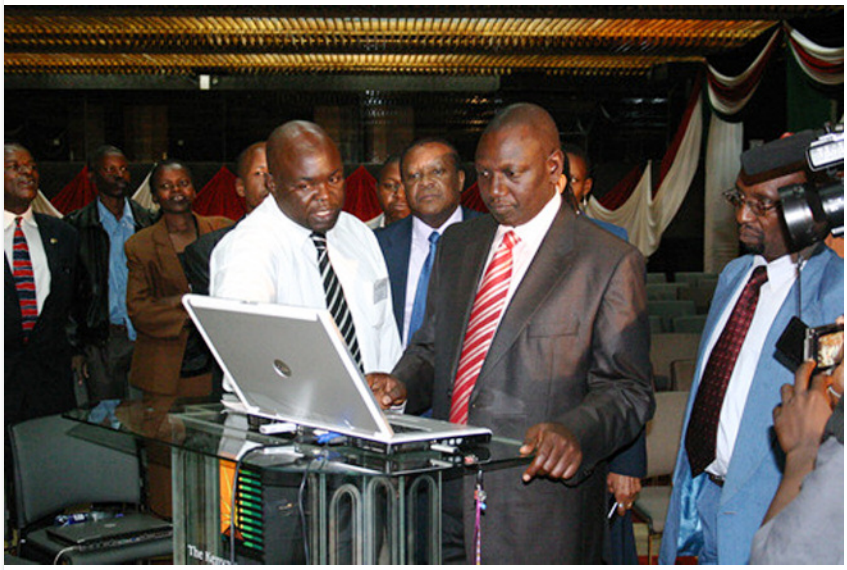
- creating awareness, ensuring commitment, and assessing needs for capacity development
- strategy development
- strengthening institutional capacities through training in information communication management (ICM) and provision of equipment
- completion of work from Phases 1 to 3, building of institutional repositories and
- development of the national repository of Agricultural Science and Technology, plus development of a marketing strategy and initiation of promotion to a wider stakeholder group

KAINet was officially launched in May 2009 at a ceremony presided over by the Minister of Agriculture, Hon. William Ruto.

"This KAINet partnership is creating the synergy that will foster the free flow of information from its generation, harvesting, proper storage, and sharing, to make sure the information is used to better the quality of our agriculture and the livelihoods that are derived from agriculture. This network will be a tool to enhance the quality of our policy decisions and our policy initiatives."

— Hon. William Ruto, Minister of Agriculture, speaking at the official launch of KAINet

Figure 2 – Hon. William Ruto, Minister of Agriculture of Kenya at the KAINet launch



Planning and Advocacy – Initial activities focused on consultations between national and international stakeholders in agricultural research, education, and extension to establish the basis for the network community and its activities for the first phase of three years. Participatory consensus-building workshops were held to raise awareness of the initiative and gain commitment at all levels throughout the five institutions including management, researchers, and information and communication specialists. Capacity assessments were also completed.

Capacity Development – Capacity development activities addressed three dimensions, namely the enabling policy level, organizations, and individuals. Firstly, an overall KAINet strategy and implementation plan was finalized with inputs from a broader stakeholder community and was formally presented during the official launch of KAINet in May 2009. In the organizational dimension, the lack of functional institutional strategies and policies in research information sharing and communication to their development through write-shops with institutional teams followed by peer review sessions with partners from other institutions. Individual training to imbue awareness, understanding and technical skills in institutional planning, information management and communication was provided to 55 staff from the KAINet member institutions. Training was delivered in a variety of formats in areas ranging from policy, planning and management, marketing of information products and services, through to technical aspects of information systems management.

The training was also valuable for professional development and staff recognition from management of their institutions, who had included KAINet project activities in staff performance contracts. Other benefits included a network of professional colleagues with whom they were able to share experiences and challenges related to their work, a positive change in the perceived negative attitude of the users of the information services towards library information services, and improved visibility of scientists who gradually came to appreciate the benefits of sharing their outputs.

KARI's IT Systems Administrator was trained by FAO in trouble-shooting software tools and methodologies, and provided technical advice across the network to address technological problems experienced at the different institutions. In consultation with project partners, he led the development of guidelines for validating information for input into the repositories, establishing the central KAINet repository and the KAINet website, and served as a link for the developers of the technologies in FAO. Crucial IT equipment was also provided to some of the participating institutions, to upgrade facilities to a more or less common standard.

Development of repositories – Existing databases and repositories were reviewed and upgraded into the institutional repositories, and where necessary content was harvested from global databases where it was no longer available locally. In all the five participating institutions, targets for content development were surpassed, with enhancements made to workflows, information management tools and methodologies, which enabled systematic capture of full text documents and metadata including historical digitized documents and digitally-born ones.

Promotion of KAINet – Promotion of the network started immediately after its inception to create awareness through various internal events in the member institutions, articles written in institutional publications and presentations at conferences and workshops. A more formal marketing strategy was developed for KAINet, focusing in particular on its flagship product the KAINet e-repository.

Progress and Achievements

- ***Establishment of KAINet:*** The Kenya Agricultural Information Network (KAINet) has been formally established with the endorsement of the Minister of Agriculture (Figure 2), together with its strategy and implementation plan. KAINet is registered at national and sub-regional levels as a non-profit making Trust to provide a forum for participation by all stakeholders, with a national stakeholders' forum, a board of trustees and a network management committee.
- ***Enabling Policies and Stengthened Capacities:*** KAINet is integrated into the national and institutional policies and strategies and its outputs and resources such as the institutional and national repositories of agricultural information will complement on-going, national and global initiatives aimed at promoting sharing of information. and support development of human capacity in ICM through training programmes.
- ***Repositories of research outputs:*** The national repository of Agricultural Science and Technology Information is accessible through the KAINet website: www.kainet.or.ke. In addition, the five KAINet member institutions have their own instititional repositories accessible on institutional LANs and some of them are published on the Internet (e.g. KARI). The repositories include over 35,000 metadata records generated by the institutions that conform to international coherence standards to facilitate access and sharing, and about 1,500 full-text documents. In addition, three of the five centres are successfully exporting content to FAO's global public domain AGRIS database, thereby further increasing the international accessibility of Kenyan research outputs (Figure 3).

Financial sustainability of KAINet is a priority. To ensure this, KAINet activities have been included in institutional budgets of pilot centres and in performance contracts of ICM staff.

What's next – Following the success of the first phase of KAINet development, the principal institutional actors that founded the network will need to address the following areas:

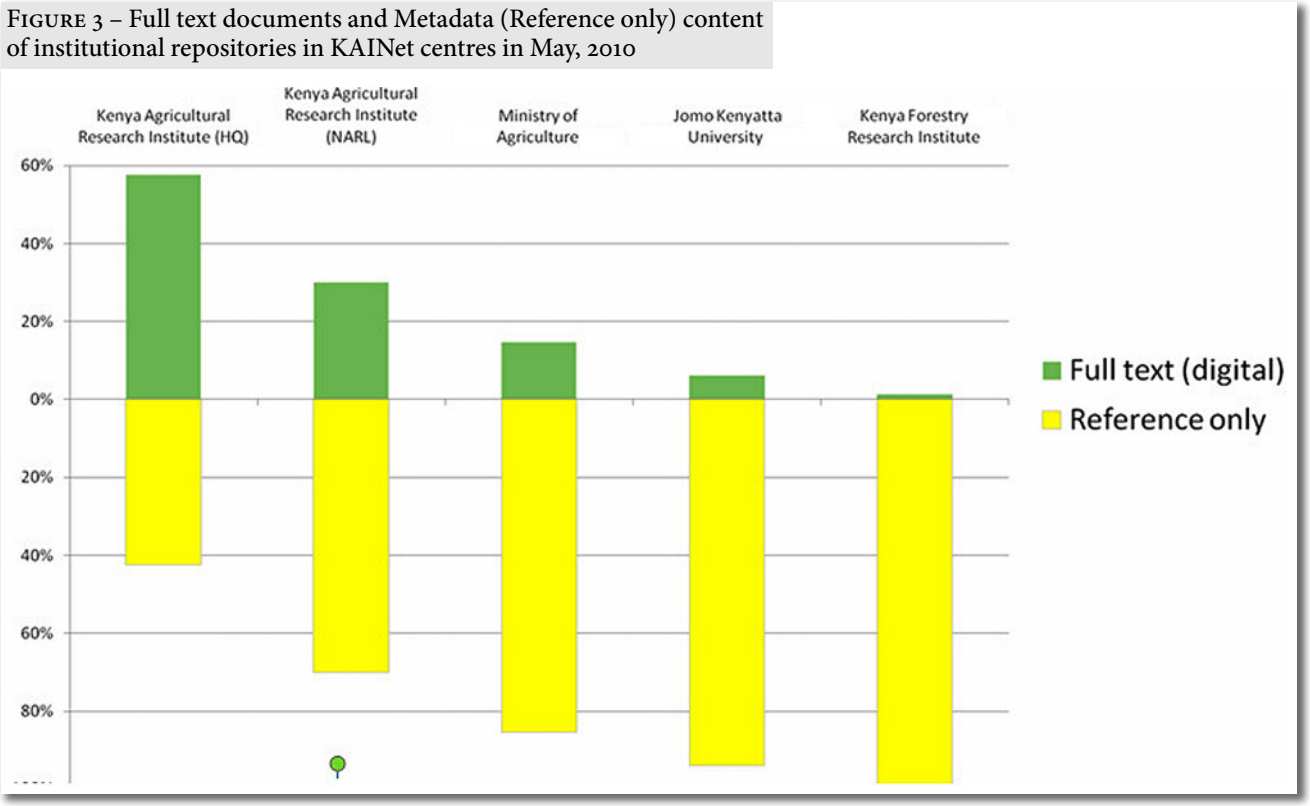
- develop incentives for researchers and academics to place more of their research outputs in the institutional repositories;
- build an evidence base of the benefits of the institutional repositories in increasing the accessibility of research outputs by analysing their use over time;
- enhance communication activities and advocacy materials aimed at increasing awareness of the KAINet resources and attracting contributions;
- consolidate membership by attracting wider participation from agricultural-sector stakeholders in research, education, and extension sectors in Kenya.

Access to the outputs of public research in agriculture: Ghana

The Actors

This initiative focused on the pilot implementation of an electronic repository as part of the Ghana Agricultural

Figure 3 – Full text documents and Metadata (Reference only) content of institutional repositories in KAINet centres in May, 2010



Information Network System (GAINS). The Ghana AGRIS Pilot Project (GAPP) was established in 2007 to support the efforts and activities of the Ghana Agricultural Information Network System (GAINS) which has been in existence since 1992. GAPP was to promote the philosophy of the open access to information approach espoused by the CIARD initiative. The main goal of GAPP was to develop open access to public domain scientific and technical information on agriculture.

Agricultural research and development are influenced by unimpeded flow of information among the sector's stakeholders including lecturers, researchers, students, policy makers and farmers, to name a few. In the past, however, provision of agricultural information support services in Ghana remained largely uncoordinated and several useful documents were scattered in various agricultural institutions and among researchers and generally had limited distribution. Besides, useful international and local journal articles were also generally not accessible to Ghanaian lecturers and researchers. This situation was partly remedied in 1991 when the Ghana Agricultural Information Network System (GAINS), a network of 18 Ghanaian agricultural and academic libraries, was established to revive the library and information system in the agriculture sector to support agricultural research and development. GAINS, therefore, plays a frontal coordinating role in the harnessing and sharing of agricultural information (both locally and internationally generated) in Ghana. Though GAINS has performed well over the years and has considerably increased and facilitated the flow of information among stakeholders, it faces some operational challenges not the least of which are inadequate capacity to harness and share generated agricultural information especially in digital forms to be readily accessed by potential end users. It is against this background that the GAPP was established in 2007.

The GAPP was a partnership involving seven national pilot institutions (PIs). They were the Cocoa Research Institution of Ghana (CRIG), the College of Agriculture Education, Mampong-Ashanti (CAGRIC), the Animal Research Institute (ARI), the Food Research Institute (FRI), the Forestry Research Institute of Ghana (FORIG), the Institute of Scientific & Technical Information (INSTI) and the Ministry of Food and Agriculture (MoFA). The Project was managed by a Project Management Committee (PMC) consisting of the Director of INSTI and a member each from the seven PIs. INSTI coordinated the Project and also housed its secretariat. These institutions were the Pilot Institutions of the GAPP and each was represented by one member on the PMC which directs the project. It was therefore the responsibility of each Pilot Institution to nurture and actively prosecute GAPP's agenda at the institutional level and also see to the project's overall success at the national level. Periodic consultations and meetings (official and unofficial) ensured a sustained sense of joint ownership of the project.

Three main categories of stakeholders stood to benefit from the GAPP. The first category consists of the staff at the Pilot Institutions. The institutional benefit is largely from the higher national and international visibility to be gained for becoming a participant of the Project and the material and human resource development support it would attract from the Project. The second beneficiaries of the GAPP are the staff of the PIs who stand to benefit greatly from the easy and rapid access to scientific literature which otherwise would be beyond their reach. In addition, the staff would be able to put their research in the public domain for easy access. The attendant benefits of this are many, not the least of which is the potential for interaction and collaboration with counterpart researchers elsewhere. The third category of beneficiaries is made up of students whose research would be much facilitated by the easy access to pertinent literature. With proper and sustained national awareness creation on the GAPP, other categories of stakeholders such as non-governmental organizations, farmer-based organizations, agro-based companies, individual farmers etc. would become beneficiaries of the project.

The Activities

The work was managed in three phases, with work plans agreed by the PMC derived from a master work plan from the GAPP proposal document to address the six specific objectives.

Phase one (October 2007 to March 2008) – The GAPP was designed to have institutional ownership, and 15 institutions of the GAINS with potential to become partner institutions of the GAPP were visited by a team from the INSTI to discuss the GAPP document, instill a sense of ownership of the project among the institutional management and ICT/M staff and select some of the institutions to participate in the project. Six of these institutions were finally selected to participate in the Project based on potential capacity to deliver, and MoFA was added because of its pivotal position as a key governmental stakeholder. Project Management and Institutional Management Committees were established to lead the Project at national and institutional levels respectively. Awareness creation, consensus and confidence building workshops activities were aimed at key actors in the project. Subsequently, Members of the PMC were trained by the Pan African Institute for Leadership and Governance Studies, Accra, on project planning, execution, monitoring, and evaluation.

Phase 2 (April 2008 to March 2009) – There were six areas of activity:

- establishment of electronic repositories on public domain scientific and technological information on agricultural science and technology at partner institutions
- finalization of ICT/M and Workflow policy documents
- knowledge sharing within the GAINS network
- enhancing capacity (human and material)

- improving collaboration among network institutions
- development of a national portal for agricultural sciences and technology information

Phase 3 (April 2009 to March 2010) – This phase was devoted to the establishment of the national portal, establishment of e-repositories centrally and at the separate institutions with full texts and metadata on agricultural sciences and technology information, knowledge sharing of and sensitization on agricultural and technology information at the national level.

Progress and Achievements

- ***Awareness of GAINS/GAPP:*** Considerable awareness was created among the staff and management of the Pilot Institutions for some who initially found the project a novelty and difficult to fathom. Issues such as: copyright, institutional repositories and open access to public domain agricultural scientific and technical information are now fairly well understood. With awareness created, many of the staff became participants of the project, which enhanced sustainability of the project and a report was published.
- ***ICT/M Policies and Workflows:*** The pilot institutions developed their ICM policy and strategy documents. Further, all CSIR research institutes benefited from this activity in the sense that the CSIR decided that the ICM policy and strategy document which was being developed for ARI, FORIG, FRI and INSTI should be extended to all CSIR research institutes. A policy document for the GAINS was also launched. The workflows for various information documents generated within the Pilot Institutions were also revised. They incorporated the capture of documents for the institutional repositories.
- ***Knowledge Sharing Framework:*** A GAINS Knowledge Sharing Framework was developed and validated for sharing agricultural information in the GAINS network on ICT/M and strategies and policies. This framework will greatly facilitate the exchange of agricultural, scientific and technological information and knowledge among the GAINS member institutions and as a result strengthen GAINS.
- ***Institutional ICT Capacity:*** Some basic ICT equipment including PCs, scanners, UPS, printers and external hard drives were supplied to Pilot institutions. A server was also purchased for the GAINS coordinating Centre. Key staff of the PIs had training on marketing of agricultural information services, management of electronic documents and AGRIS tools and methodologies, website development and management, and on copyright and institutional repositories. Capacities of the PIs were therefore greatly enhanced for GAPP's work.
- ***Institutional Repositories:*** Institutional e-repositories were established at almost all the Pilot Institutions, albeit at varying degrees of scale. Metadata and associated full-text documents are being made available to their staff on institutional intranets and later will be accessible on the Internet to the public. Contents of the various institutional repositories in March 2010 are presented in Figure 4.
- ***National Portal:*** Although not launched by the end of the project, specifications for the national portal for

Figure 4 – Status of document holdings in GAINS centres in March 2010

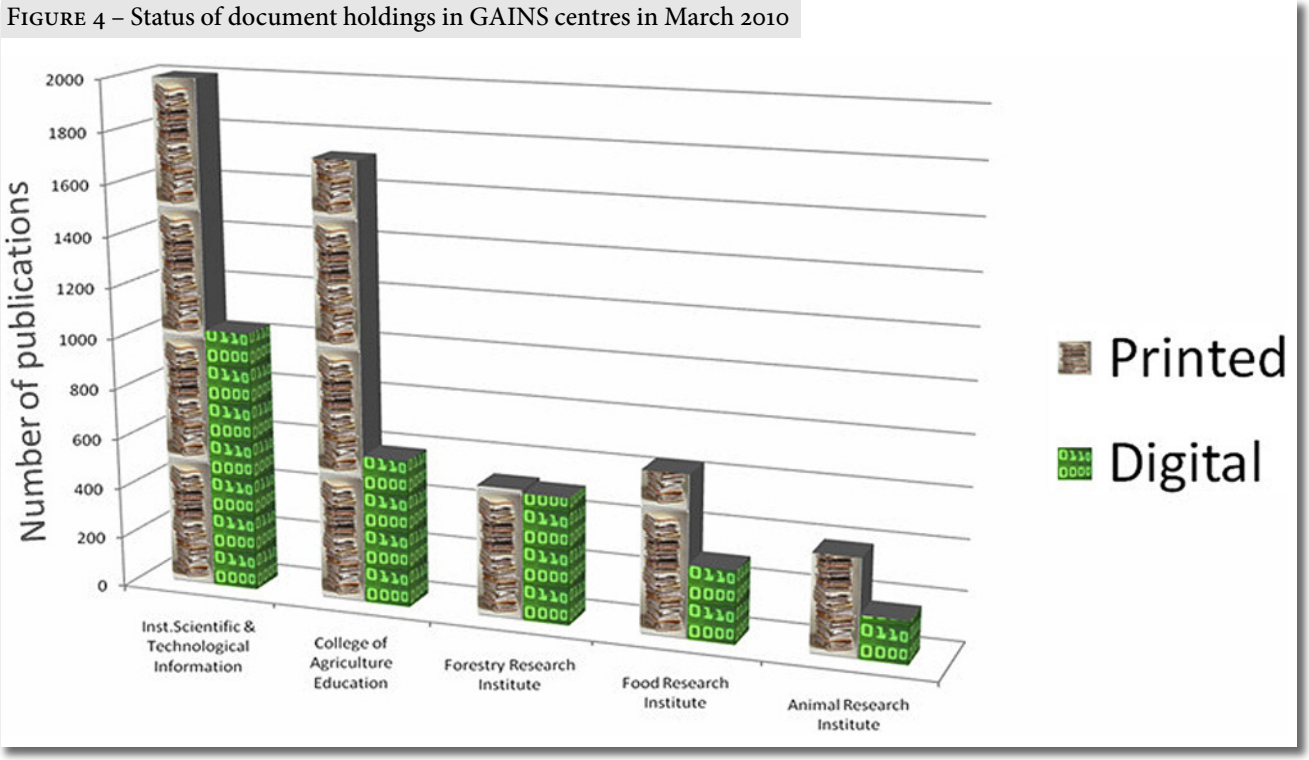


Figure 5 – Key enabling factors and lessons learned for similar initiatives

- Basing the project on national **policies and strategies** in both the agricultural and information and communication sectors was critical for the project's success and is expected to contribute to its sustainability.
- Working in **partnerships**, both national and international, helped to deliver joint planning, rationalizing of resources, sharing of skills and experiences, and cross fertilization of ideas.
- **Piloting** the network with a limited number of national institutions allowed the partners to learn and devise workable solutions before expanding the network.
- The management/steering committees played important roles in involving the **management** of institutions, promoting the network and guiding project activities.
- Linking the project to **institutional priorities/plans** added credibility, ensuring that it would enhance existing work, rather than leaving it as a standalone initiative.
- **Phasing** of activities provided a systematic approach. In particular, the initial phase of planning and building partnerships was critical for the success of the project as it provided an understanding of the institutions' ICM needs and the basis for collaboration.
- **Networking** contacts provided the basis of project operations. Telephone and e-mail groups were essential for constant communication among partners and facilitated sharing of experiences.
- Use of **participatory approaches** such as "write-shops" in developing ICM strategies, and stakeholder consultations in the institutions, were very effective in ensuring broad representation of stakeholders in strategy development and in imparting strategic planning skills.
- Development of adequate **ICM capacities** (including equipment) was essential for development of open repositories, and these capacities should preferably be built in the early phases of the project.

agricultural sciences and technology information had been agreed upon by the GAINS stakeholders and work had began on developing the portal. The portal will be key in facilitating access of the GAINS' central e-repository into which metadata from the Pilot Institutions can be placed.

What's Next – The GAPP was established in 2007 with the main objective of establishing electronic repositories at selected Pilot Institutions. The achievements are substantially in line with the objectives. Given the success of the GAPP, the principal institutional actors now need to address the following areas:

- Facilities: Pilot institutions should provide adequate equipment for ICM activities in the institutions to cope with the volume of work anticipated in the future. Institutions should also invest in the security and protection of the ICT facilities from computer viruses.
- Incentives: The institutions will need to develop policies and/or schemes to ensure that for researchers and academics place their research outputs in institutional repositories.
- Involvement of more institutions: GAINS should extend the activities initiated by the Pilot Project to all its member institutions involved in the generation, management and dissemination of agricultural technological and scientific information in Ghana. For example, a number of agricultural faculties and institutions exist in Ghana with staff and students generating much-needed information.
- Financial Sustainability: the Pilot Institution's Management will need to sustain financial allocation for ICT/M development and the necessary ICT/M infrastructure acquisition and maintenance to supplement technology provided by the project.

Lessons learned

Strong progress has been made in the pilot national networks of institutional repositories of agricultural research outputs in Ghana and Kenya, with enabling institutional policies formulated through participatory approaches and significant quantities of full text content now accessible in digital format through organizational repositories (www.gains.org.gh and www.kainet.or.ke). The experiences in Kenya and Ghana were carefully documented in the form of case studies summarized in Figure 5, completed in collaboration with the relevant sub-regional organizations of ASARECA and West and Central African Council for Agricultural Research and Development (CORAF/WECARD).

Although the two case studies were initiated before the conceptualization of the CIARD framework, first conceived in 2008, these lessons learned from Kenya and Ghana were fed into the consultation process for the CIARD initiative in 2009, in terms of the formulation of the Checklist of Good Practices and the detail of the Pathways.

Apart from the successful outcomes, these two interventions faced and continue to face a range of challenges. Those include shortage of skilled human resources, inad-

equate technology resources, poor Internet connections, and unreliable electrical power supplies; and these barriers have to be contended with for the foreseeable future. The most serious barrier to further development of the open archives in Kenya and Ghana is the reluctance of scientists to archive the outputs of their research in institutional repositories, and this barrier is known to be a worldwide issue that by no means is confined to developing countries. The work to date did not study the underlying reasons for this reluctance, and this remains a serious challenge that has to be addressed.

Contact Information

Justin Chisenga
Food and Agriculture Organization
Accra, Ghana
Justin.Chisenga@fao.org

Richard Kedemi
Kenya Agricultural Research Institute
Kenya
rkedemi@kari.org

Joel Sam
Institute of Scientific & Technical Information
Ghana
jsam@insti.csir.org.gh

Stephen Rudgard
Food and Agriculture Organization
Rome, Italy
Stephen.Rudgard@fao.org

Franz Martin
Food and Agriculture Organization
Rome, Italy
Franz.Martin@fao.org

The CIARD RING, an Infrastructure for Interoperability of Agricultural Research Information Services

Valeria Pesce, Ajit Maru and Johannes Keizer

Editor's Note: Paper presented at the IAALD XIIIth World Congress, Scientific and Technical Information and Rural Development, Montpellier, 26–29 April 2010.

Abstract: The Coherence in Information for Agricultural Research for Development (CIARD) Routemap to Information Nodes and Gateways (RING) is a portal offering an interlinked registry of existing information services in agriculture that covers both information services and sources. In the RING, the definition of "service" includes any form of providing information from one server instance (website, mail server, web services, XML archive…) to many clients (browsers, email clients, news readers, harvesters…) The services registered in the RING are described in detail and categorized according to criteria that are relevant to the use of the service and its interoperability. The RING categorizes and interlinks the featured services according to criteria such as: standards adopted, vocabulary used, technology used, protocols implemented, level of interoperability etc. In addition, it features detailed instructions on how the registered services can be "interoperated". The vision is that the RING will become the common global technical platform for the community of agricultural information professionals for accessing, sharing and exchanging information through web services. This paper describes how the RING provides an infrastructure for enhancing interoperability of information sources and thus paves the way towards better accessibility of information through value-added and better targeted services.

Resumé: La Carte routière des pôles et portails d'accès à l'information (RING) de la Cohérence de l'information sur la recherche agricole pour le développement (CIARD) est un portique qui offre un registre interconnecté des services actuels d'information agricole qui couvre aussi bien les services d'information que les sources. Dans le RING, la définition de « service » inclut n'importe quelle forme de fourniture d'informations, d'un serveur quelconque (le site Web, le serveur de courriel, les services de toile, l'archive de XML…) à une multitude de clients (les navigateurs, clients e-mail, les lecteurs de nouvelles, les récolteurs…) Les services enregistrés auprès du RING sont décrits en détail et classés selon les critères qui sont pertinents à l'usage du service et de son interopérabilité. Le RING classe et lie les services présentés selon des critères comme : les normes adoptées, le vocabulaire utilisé, la technologie utilisée, les protocoles exécutés, le niveau d'interopérabilité, etc. Par ailleurs, il présente des instructions détaillées sur comment les services inscrits peuvent être « interopérables ». Le RING a pour vision de devenir la plateforme technique mondiale commune à la communauté des professionnels de l'information agricole pour l'accès, le partage et l'échange d'informations par le biais des services Web. Cet article décrit comment le RING fournit une infrastructure pour améliorer l'interopérabilité des sources d'information, et pave ainsi le chemin vers un meilleur accès à l'information grâce à la valeur ajoutée et des services mieux visés.

Resumen: El Mapa de Ruta a los Nodos de Información y Medios de Acceso (RING, sus siglas en inglés) para Coherencia en la Información para la Investigación Agrícola para el Desarrollo (CIARD, sus siglas en inglés) es un portal que ofrece un registro entrelazado de servicios existentes de información en agricultura que cubre tanto los servicios como las fuentes de información. En RING, la definición de "servicio" incluye cualquier forma de suministro de información desde un servidor (sitio web, servidor de correo, servicio Web, archivo XML…) a muchos clientes (navegadores, clientes de correo electrónico, lectores de noticias, cosechadoras…). Se describen detalladamente los servicios registrados en RING y se clasifican según criterios que son pertinentes para el uso del servicio y su interoperabilidad. RING clasifica e interconecta los servicios presentados según criterios como normas adoptadas, vocabulario utilizado, tecnología utilizada, protocolos implementados, nivel de interoperabilidad, etc. Presenta, además, instrucciones detalladas sobre cómo los servicios registrados pueden ser "interoperables". La visión es que RING se convierta en la plataforma técnica común global para la comunidad de profesionales en información agrícola para tener acceso, intercambiar y compartir información mediante servicios Web. Este trabajo describe cómo RING ofrece una infraestructura para mejorar la interoperabilidad de las fuentes de información y, por tanto, prepara el camino para una mejor accesibilidad de la información mediante servicios de valor agregado que estén mejor direccionados.

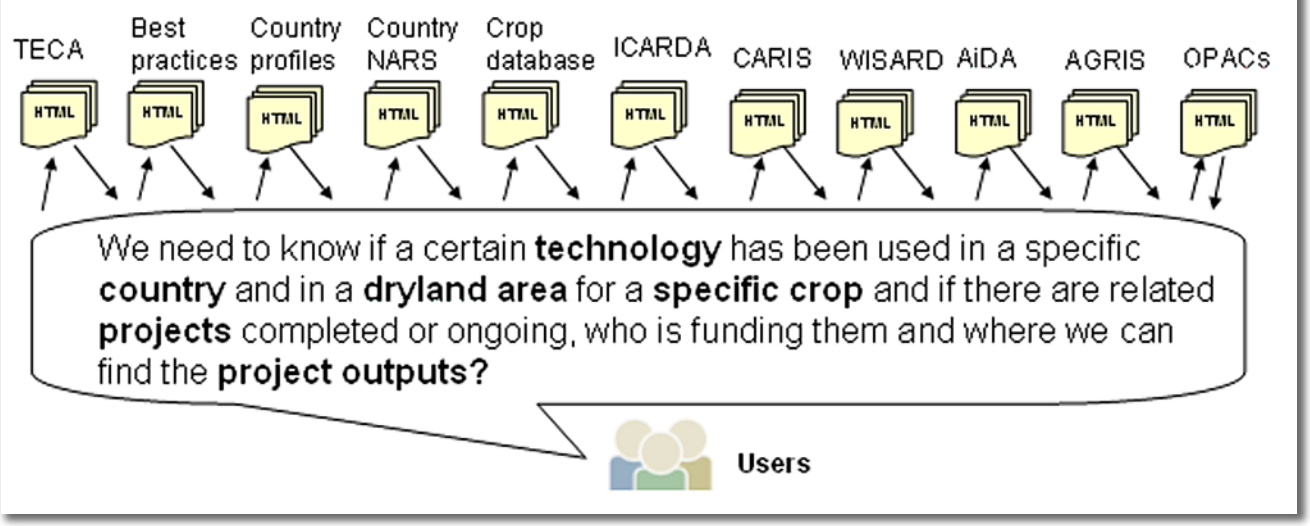
Background and Rationale

While information needs for agricultural research, innovation and development become more and more complex, agriculture-related information is widely distributed across sources that in most cases do not share and exchange data among themselves. On the whole, agricultural information on the web is not organized for easy access. The consequence is that those who need it do not find it or at least do not find all the relevant pieces of information that are potentially available to meet their specific needs.

Users cannot be aware of all the available sources of information, as there is no comprehensive list or directory of existing information services in agricultural research for development (ARD). Even out of the sources of which they are aware, they have to search several, which are not and cannot be comprehensive and therefore provide partial answers to complex questions. Lastly, they cannot always get the information in a format or through a channel that is accessible to them, be it due to language barriers, literacy barriers, technological or infrastructural barriers (Figure 1).

The problem is not the quantity of existing information

Figure 1 – Information Needs and Information Services in Recent Year



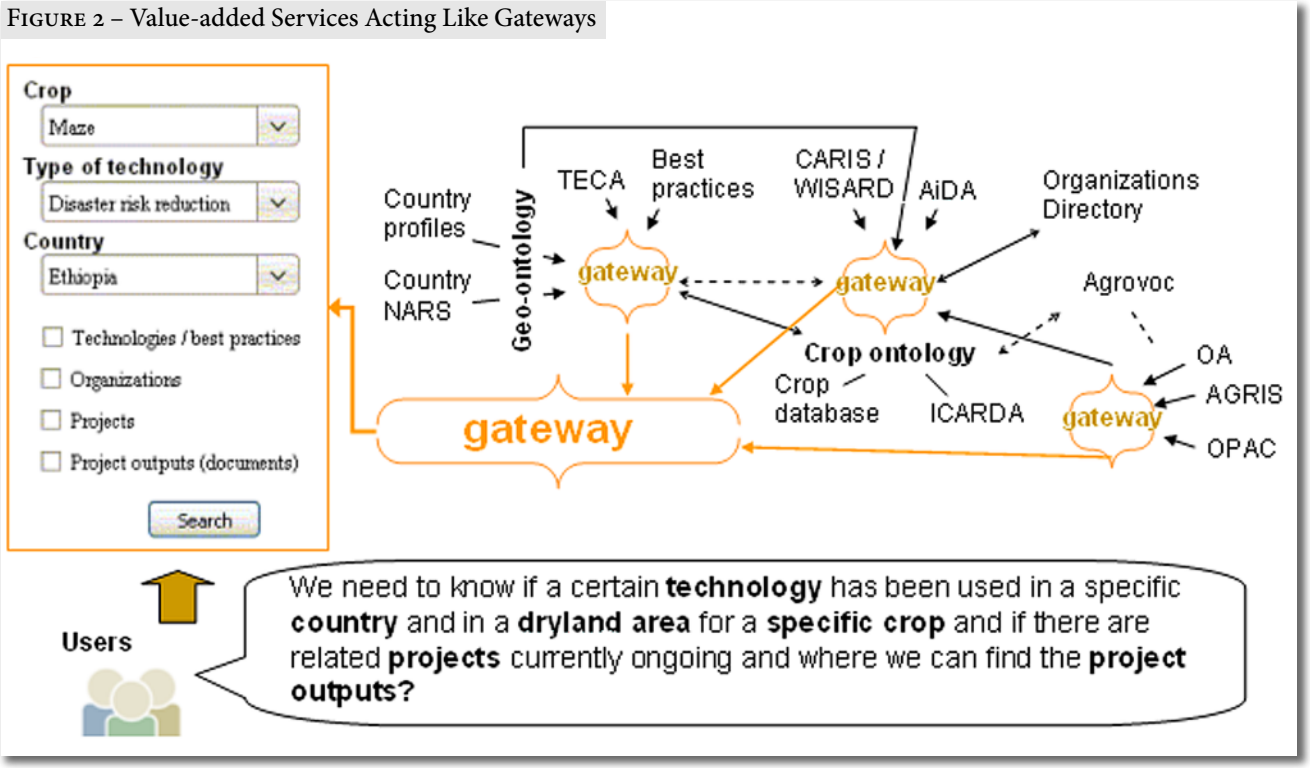
sources or that they are different but sometimes overlapping in coverage or that none of them is comprehensive: on the contrary, the fact that information sources are managed at the "lowest" level possible (from a geographic or an administrative point of view) in a distributed way allows for better and more sustainable maintenance.

However, from the consumer's point of view, this makes valuable information out of their reach, keeping it hidden and therefore only virtually available.

How can users find the information they need when it is residing in widely scattered distributed sources? Only through highly user-targeted services that enable their specific audience to search, collate and integrate information from various sources acting as gateways to them. Such value-added services re-package the collated information and make it available through different browsing and search options, different formats, and different channels according to the target users.

Creating such value-added information services in agriculture remains a major challenge, as they should by definition interface several information sources that use different forms of semantic organization, different languages and different levels of aggregation, and then integrate, semantically enhance and re-package the original information. Therefore value added services cannot be built without at least an awareness of what others have

Figure 2 – Value-added Services Acting Like Gateways



Programs Focusing on Information Sharing and Interoperability — 2000–2008

- **GFAR** (Global Forum for Agricultural Research) – *http://www.egfar.org/egfar/website?contentId=-1&*
- **Global RAIS** (Regional Agricultural Information Systems) – *http://www.fao.org/docs/eims/upload/216230/GLOBAL.RAIS_Proposal.pdf*
- **ICM4ARD** Programmes (Information Communication Management for Agricultural Research and Development) – *http://www.fao.org/docs/eims/upload/210060/GlobAL-RAIS_Inter-regional_Workshop_2004_Rome.pdf*
- **FAO** (Food and Agricultural Organization of the United Nations) – *http://www.fao.org/*
- **AGRIS** initiative – *http://agris.fao.org/*
- **Consultations on Agricultural Information Management** –
  2000 – *ftp://ftp.fao.org/unfao/bodies/coaim/coaim-2/AC502e.doc*
  2002 – *http://www.fao.org/docrep/meeting/005/y7963e/Y7963e00.htm#Top*
- **IICD** (International Institute for Communication and Development) – *http://www.iicd.org/*
- **DFID** (Department for International Development - United Kingdom – *http://www.dfid.gov.uk/*
- ***Fertile Ground Study*** – *http://www.ciard.net/history*
- **CGIAR** (Consultative Group on International Agricultural Research) – *http://www.cgiar.org/*
- **ICT-KM** (Information and Communication Technology-Knowledge Management) Programme – *http://ictkm.cgiar.org/*
- **IISAST** (International Information Systems for Agricultural Science and Technology) global consultations in 2005 and 2007 – *http://www.ciard.net/history*

done: which sources are available, how to tap into them, how to exploit their semantics (Figure 2).

In order for this to happen, information sources need to become more easily "discoverable" and to allow other services to semantically query, re-use and re-package their information; in other words they must become truly interoperable and allow for easy automatic retrieval of information, while work on mapping between vocabularies or advanced natural language processing must be done to improve the semantic accessibility of information.

The community of agricultural information specialists has been well aware of these issues for some time and the issue of information sharing and interoperability has been at the core of programmes carried on by international organizations in the years 2000-2008 and in 2008 several major international organizations have joined forces under the **C**oherence in **I**nformation for **A**gricultural **R**esearch for **D**evelopment (CIARD) initiative.

Within CIARD, the Global Forum on Agricultural Research (GFAR), building on the former idea of the GLOBAL.RAIS "Web Ring" that aimed at improving integration of national and regional agricultural information systems, proposed and implemented the "CIARD RING": the **R**outemap to **I**nformation **N**odes and **G**ateways (RING), a registry of information services in agriculture that is to act as a route map towards better interoperability of information services.

This paper describes how the RING provides an infrastructure for enhancing interoperability of information sources and thus paves the way towards better accessibility of information through value-added and better targeted services.

The RING As a Registry and As a Route Map

The CIARD RING is a web-based registry of existing information sources in agriculture research and development that supports CIARD in monitoring and improving the accessibility of information by:

- providing a map of where available sources of information (on a certain subject domain, in a certain format, for a certain audience) can be found and instructions on how they can be effectively searched;
- providing the examples of existing services that represent good practices on how interoperability can be implemented;
- making the level and modes of interoperability of information services more explicit;
- providing all the necessary information for building value-added integrated services that re-package information and make it accessible in different ways.

The RING is first of all a registry where detailed information about information sources is collected through a submission form and made searchable through a search engine.

It is important to define what an "information source" is and why in the context of the RING the terms information source and information service are used interchangeably: in today's information architectures, the distinction between the two is fluid. This is why the concept of "nodes and gateways" in the RING covers both information sources and information services, that is, both "static" files available in some structured format (like XML or RDF, but also data text files like .csv) and

interactive services like search engines and web services. The reason is that both ways of making information available can be made interoperable and can contribute to improve the accessibility of information on the whole.

In a broader sense, the definition of "service" in this context includes any form of providing information from one server instance (website, mail server, web services, XML archive…) to many clients (browsers, email clients, news readers, parsers, harvesters…).

Examples are:

- services that provide RSS feeds
- services that provide XML exports of information based on agreed metadata sets (e.g. the AGRIS data providers)
- Open Archive Initiative (OAI) data providers
- OAI harvesters
- services that offer web services for accessing and re-using their information
- RDF stores
- SPARQL engines

The services registered in the RING are described in detail and categorized according to criteria that are relevant to the use of the service and its interoperability. The RING categorizes and interlinks the featured services according to criteria such as: standards adopted, vocabulary used, technology used, protocols implemented, level of interoperability etc. In addition, the RING features detailed instructions on how the registered services can be "interoperated" (Figure 3).

This is why the RING goes beyond being a registry and acts in fact as a "route map":

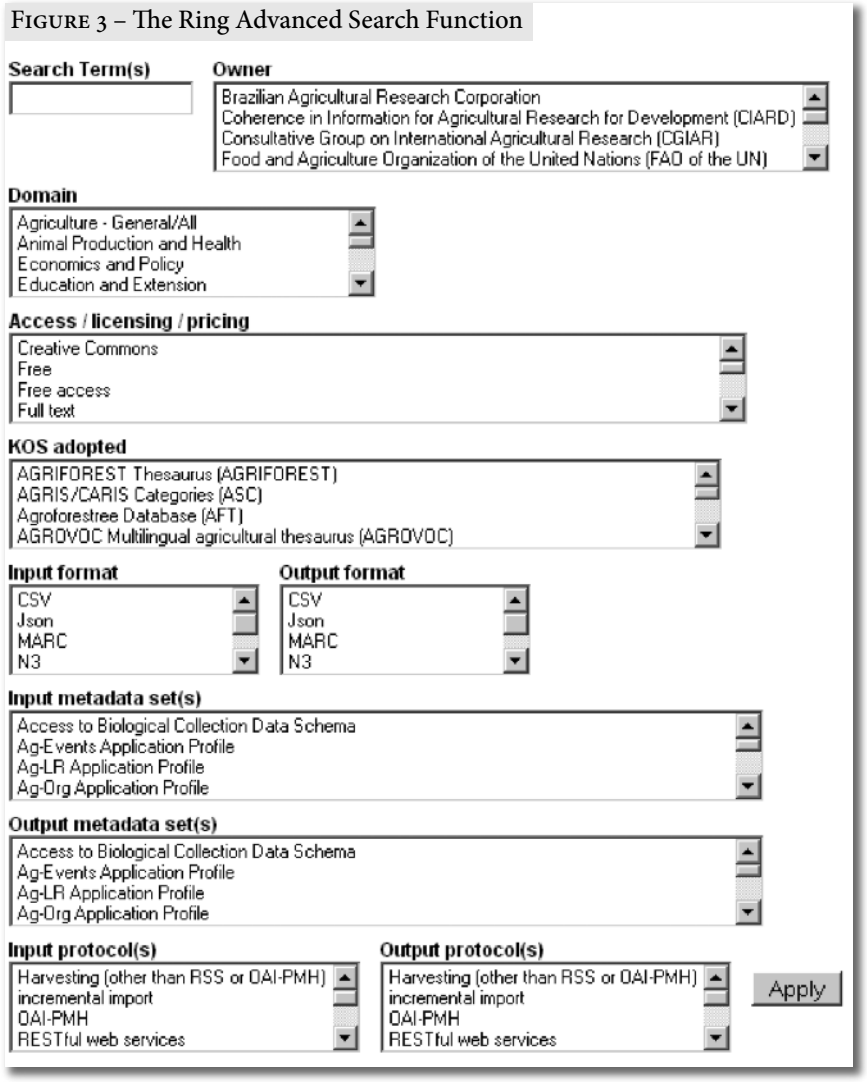
- Thanks to the way services are indexed and made searchable, users are easily guided to find the type of information service they are looking for.
- Thanks to the way services are indexed and described, users can learn how to "exploit" them for building value-added integrated services.
- Thanks to the type of information gathered, a map of existing services can be drawn, highlighting the relationships between them and the flow of information among them.

The potential impact of the RING is not so much in the collected information itself as in what can be built out of it. Providing structured information on the metadata sets, the formats, the protocols and the vocabularies used in each registered source will facilitate the building of applications like:

- services that offer a common browsing or searching interface to different sources;
- aggregating and harvesting services;
- integrated services providing relations between entities (organizations, projects, experts, documents) through semantic-web technologies;
- services that re-package information and make it available through different channels (text messaging, radio etc.);
- services that interface the different knowledge organization systems (KOS) used by different sources;
- applications providing value-added services like digests, bibliographies, best practices, surveys etc.

While the final objective is that of benefitting the consumers of agricultural information by increasing the quantity and improving the quality of value-added targeted services, the immediate users of the RING are mainly information professionals and web developers, who can on

Figure 3 – The Ring Advanced Search Function



the one hand provide the relevant technical information about their services at registration time and on the other exploit the technical information shared by others in the RING in order to tap into the existing sources and feed their own systems.

Also the managers of information services may find the RING useful in different ways: they can identify useful information sources for the services they manage, they can find information services into which they can feed their contents for better dissemination and they can publicize the services they manage by just registering them.

Finally, once the RING is populated with a significant number of information services, also consumers of agricultural information may find it useful as a "bookmark" list to relevant information services in agriculture.

FIGURE 4 – A geographic map of the physical location of the registered service.

The Present Stage and the Way Forward

The RING is available at http://ring.ciard.net (Figure 4) and is currently in its initial phase.

The first phase, just started, consists in building the registry.

In this phase, the RING is gathering information on which information sources / services are currently available and how to tap into them.

The metadata that are collected to allow:

- categorize and interlink the registered services according to specific criteria: standards adopted, vocabulary used, technology used, protocols implemented, level of interoperability, target audience etc.;
- link the services to each other through provider/consumer relationships;
- feature detailed instructions on how the featured services can be "interoperated".

The services featured in the RING are submitted directly by their managers and technical staff, which ensures ownership and reliability of the data. Any person who is responsible for an information service can register it. Each record describing a service must link to the record of an organization / institution registered in the system: these records can be created on the fly while registering the service or can be just referenced if they exist. A mandatory element is the email address of the institution: in order to ensure the correctness of attribution of the services to their owners, the organizations responsible for the service will be alerted upon submission and periodical checks will be run by the RING administrators.

Already in this first phase information professionals and web developers can exploit information in the RING to build advanced services that tap into the registered services.

The second phase will start when the technical information collected about the services is detailed and structured enough and when the number of registered services that have a good level of interoperability is significant enough: at this stage, some advanced services can be built semi-automatically directly on the RING website.

Examples are:

- an Open Archive Initiative (OAI) harvester harvesting all the registered OAI providers;
- an RDF viewer/navigator of the registered RDF stores;
- sample thematic RSS aggregators that harvest from the registered RSS feeds;
- sample consumers of web services.

Additionally, sample programming code on how to implement similar services will be provided on the RING website.

Conclusions

The effectiveness of the RING in achieving its objectives will be proportional to the number and quality of the services registered in the system. A high number of interoperable services correctly indexed in the system will allow to provide a comprehensive registry and a real infrastructure that can be leveraged for building new services; the examples of advanced services featured in the system will provide guidance and incentives to the building of new ones; and the collected data will make it possible to generate interesting views on the flows of information between the services.

The core of the system and the real infrastructure for interoperability will be the registered web services: with the broad meaning of the term "web service" introduced by the advent of the so-called "RESTful" web services, these now include RSS feeds, RDF stores, SPARQL engines and OAI providers, and these are presently the best technologies for interoperability.

The vision is that the RING will become the common global technical platform for the community of agricultural information professionals for accessing, sharing and exchanging information through web services.

The RING is a CIARD project led by GFAR and its functionalities and the information therein registered are free and publicly available, thus constituting Global Public Goods that can be leveraged by any organization, person or information service.

Contact Information:

Valeria Pesce
Global Forum on Agricultural Research (GFAR)
valeria.pesce@fao.org

Ajit Maru
Global Forum on Agricultural Research (GFAR)

Johannes Keizer
Food and Agriculture Organization of the United Nations (FAO)

Keep up with IAALD!

- The IAALD Executive Committee has been active during the past year. In October of this year, Barbara Hutchinson, IAALD President began a newsletter to the membership to provide updates on the work of the organization. The newsletter is sent out over IAALD-L.
  If you want to keep up with what IAALD is doing, just watch IAALD-L.

- Not subscribed to IAALD-L?
  It is easy to do. Just go to
  http://calsmail.arizona.edu/mailman/listinfo/iaald-l

- If you are interested in
  what IAALD has done in the past year
  go to ***http://bit.ly/gxDULx*** for a general update.

International Association of Agricultural Information Specialists (IAALD) — www.iaald.org

IAALD'S MISSION is to enable its members to create, capture, access and disseminate information to achieve a more productive and sustainable use of the world's land, water, and renewable natural resources and contribute to improved livelihoods of rural communities.

To further this mission:

IAALD **CONNECTS** agricultural information specialists worldwide, providing platforms and spaces for information dissemination, exchange and knowledge sharing;

IAALD **CONVENES** agricultural information specialists worldwide, organising meetings and catalyzing dialogue among all agricultural information stakeholders;

IAALD **COMMUNICATES** and advocates the value of knowledge and information to its members and others, improving the status and practice of agricultural information management and dissemination;

IAALD **COLLABORATES** with members and other partner organisations, facilitating educational and other opportunities across agricultural information communities.

Executive Committee of IAALD

IAALD is a founding member of the International Network for Information Technology in Agriculture (INFITA) — www.infita.org

IAALD is a partner in the Coherence in Information for Agricultural Research for Development (CIARD) initiative — http://www.ciard.net

Agricultural Information Worldwide

An International Journal for Information Specialists in Agriculture, Natural Resources and the Environment

Vol. 4, No. 1, 2011

Sponsored by

In this issue:

IAALD

International Association of Agricultural Information Specialists – www.iaald.org